International Journal of Innovative Computing, Information and Control
Volume **10**, Number **6**, December **2014**

ICIC International ©2014 ISSN 1349-4198
pp. 2193–2207

NONLINEAR SET MEMBERSHIP IDENTIFICATION BY LOCALLY LINEAR EMBEDDING

Wei Chai

School of Electronic Information and Control Engineering
Beijing University of Technology
No. 100, Pingle Yuan, Chaoyang District, Beijing 100124, P. R. China
chaiwei@bjut.edu.cn

Received November 2013; revised March 2014

Abstract. *This paper proposes a novel set membership identification method for models nonlinear in their parameters. A mapping which can approximate the homeomorphism between the feasible parameter set (FPS) boundary and the $n-1$-sphere (n is the number of parameters) is constructed. First, a data set consisting of vectors uniformly sampled from the FPS boundary is mapped into a data set contained by the $n-1$-sphere. This is achieved by locally linear embedding followed by data normalization. Then, a non-parametric method based on the two data sets is used to build a mapping which approximates the homeomorphism between the FPS boundary and the $n-1$-sphere. Once this mapping is established, it can be used to map the $n-1$-sphere into an approximation of the FPS boundary. Moreover, a strategy is proposed to improve the boundary approximation. Examples show that the proposed method exhibits superior performance compared with other nonlinear set membership identification methods.*

Keywords: Set membership identification, Nonlinear models, Parameter estimation, Locally linear embedding

1. **Introduction.** System identification has attracted significant attention of the investigators and many identification methods have been developed. The classical methods, such as least squares or maximum likelihood, are based on stochastic assumptions. The noise corrupting the data is assumed to be a random variable with a known statistical property. Unfortunately, statistical properties of the noise are not available when the number of the observed data is insufficient or when the noise is of a deterministic nature.

As an alternative method, the set membership identification [1-3] considers a more realistic description of the error in the form of bounds on its instantaneous values. The purpose of this approach is to characterize the set of all the parameter vectors consistent with the data, model structure and noise bounds. This set is called the feasible parameter set (FPS). When the model is linear, the FPS is a convex polytope. However, in many applications the models are usually nonlinear in their parameters. In this case, the FPS is generally nonconvex or even nonconnected. Exact description of the FPS is an arduous task. Therefore, the key issue in nonlinear set membership identification is to find an approximate description of the FPS, which is easy to interpret and which leads to a satisfactory balance between the computation burden and precision.

The existing approaches for approximate description of the FPS in nonlinear set membership identification can be roughly classified into three categories. The first calculates a single convex set containing the FPS, such as a box [4], an ellipsoid [5,6], or a simplex [7]. This approach can give a guaranteed outer bound of the FPS, but the obtained result may be too conservative. The second gives a guaranteed inner or outer bound of the FPS with an arbitrary precision by a union of boxes [8,9]. Unfortunately, its complexity grows exponentially with the dimension of the parameter vector. Besides, it is quite difficult to construct the minimal inclusion function for complex models. The excessive conservatism of the boxes caused by the inability to find the minimal inclusion function will result in a slow convergence of the algorithm. Finally, the third characterizes the FPS by computing the points on or near its boundary [10-13]. This approach can deal with complex models for which it is difficult to construct the minimal inclusion function.

This paper proposes a novel set membership identification method for models nonlinear in their parameters. A mapping which can approximate the homeomorphism between the FPS boundary and the $n-1$-sphere (n is the number of parameters) is sought. If this mapping is established, it can be used to map the $n-1$-sphere into an approximation of the FPS boundary. To construct this mapping, the following technique is used. First, a data set consisting of vectors uniformly sampled from the FPS boundary is mapped into a data set contained by the $n-1$-sphere. This is achieved by locally linear embedding (LLE) [14,15] followed by data normalization. Then, a mapping which approximates the homeomorphism between the FPS boundary and the $n-1$-sphere is derived by a non-parametric method based on the two data sets. Besides, a strategy for improving the boundary approximation is proposed. This strategy discards the points with unsatisfied approximation errors and from them searches for new points which are on the FPS boundary.

The paper is organized as follows. In Section 2, the problem of set membership identification is formulated. The method for approximating the FPS boundary is proposed in Section 3. In Section 4, three examples are given to illustrate the performance of the method. The paper draws to a close with a section of conclusions.

2. **Problem Formulation.**

Consider a nonlinear model defined by

$$y_k = f(x_k, p) + e_k, \tag{1}$$

where $y_k \in R$ is the observable output, $x_k \in R^{n_x}$ is the regression vector, $p \in R^n$ is the parameter vector to be estimated, $f(\cdot,\cdot)$ is a vector function and $e_k \in R$ is the unobservable error. Here, $f(x_k, p)$ is nonlinear with respect to parameter vector p. It is assumed that the error e_k is bounded for each sample time k, i.e.,

$$|e_k| \le \varepsilon_k \text{ for all } k, \tag{2}$$

where ε_k is a known constant.

Suppose that a data set $\{x_k, y_k\}_{k=1}^N$ has been obtained. The set of all the parameters consistent with the model (1), the observed data set $\{x_k, y_k\}_{k=1}^N$, and the bounded error assumption (2) can be expressed as

$$P_N = \{p \in R^n : |y_k - f(x_k, p)| \le \varepsilon_k, k = 1, 2, \cdots, N\}. \tag{3}$$

The set P_N is called the feasible parameter set. Equation (3) can be rewritten as

$$P_N = \{p \in R^n : \|Y - F(p)\|_\infty^E \le 1\}, \tag{4}$$

where

$$Y = [y_1, y_2, \cdots, y_N]^{\mathrm{T}}, \tag{5}$$

$$F(p) = [f(x_1, p), f(x_2, p), \cdots, f(x_N, p)]^{\mathrm{T}}, \tag{6}$$

$$E = [\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_N]^{\mathrm{T}}. \tag{7}$$

The weighted ∞ norm $\|u\|_\infty^E$ is defined as $\|u\|_\infty^E = \max_i |u_i/\varepsilon_i|$. The boundary of the FPS P_N can be expressed as

$$B_{P_N} = \{p \in R^n : \|Y - F(p)\|_\infty^E = 1\}. \tag{8}$$

The term e_k represents the effect of modeling errors, disturbances and noise. The stochastic methods assume that the probability density function of the error e_k is (partially) known. However, the stochastic assumptions may be questionable. The statistical properties of the error e_k cannot be justified when the available data are limited. Besides, the main contribution to error e_k may be of a deterministic nature. The set membership identification considers a more realistic bounded error description (2). The purpose of this method is to characterize the FPS P_N. When the model is nonlinear, the FPS P_N is extremely complex. Exact description of the FPS P_N is an arduous task. An alternative solution is to find an approximate description of the FPS P_N. This can be achieved by three types of approaches which have been discussed in Section 1. Supposing that the FPS P_N is bounded and connected, a method for approximating the FPS boundary B_{P_N} is proposed in the next section.

3. **Set Membership Identification by Locally Linear Embedding.**

3.1. **Approximation of the FPS boundary.**

According to the theory of geometry and topology, the FPS boundary B_{P_N} is homeomorphic to an $n-1$-sphere S^{n-1}, which is the set of points in n-dimensional Euclidean space that are at a fixed distance r from a central point of that space, where r is a positive real number. A homeomorphism is a continuous function that has a continuous inverse function. Homeomorphisms are the isomorphisms in the category of topological spaces. They are the mappings that preserve all the topological properties of a given space. Two spaces with a homeomorphism between them are called homeomorphic, and from a topological viewpoint they are the same.

Therefore, the FPS boundary B_{P_N} can be transformed from the $n-1$-sphere S^{n-1}, once a homeomorphism $\varphi(\cdot)$ between them is established. However, since the shape of the FPS P_N is complex, it is not easy to find an expression for the homeomorphism $\varphi(\cdot)$. Alternatively, this paper will seek a mapping $\hat{\varphi}(\cdot)$ which can approximate the homeomorphism $\varphi(\cdot)$ between the FPS boundary B_{P_N} and the $n-1$-sphere S^{n-1}. Once the mapping $\hat{\varphi}(\cdot)$ is established, it can be used to map the $n-1$-sphere S^{n-1} into a set $\hat{B}_{P_N}$ approximating the FPS boundary B_{P_N}. Constructing the mapping $\hat{\varphi}(\cdot)$ which is illustrated by Figure 1 consists of the following three main steps:

1. Obtain a data set consisting of vectors θ_i, $i = 1, \cdots, l$, uniformly sampled from the FPS boundary B_{P_N}. Suppose that the number of the data is sufficient. Map the data set $\{\theta_i\}_{i=1}^l \subset B_{P_N}$ into a data set $\{\rho_i\}_{i=1}^l \subset R^n$ by means of LLE.
2. Map the data set $\{\rho_i\}_{i=1}^l \subset R^n$ into a data set $\{\eta_i\}_{i=1}^l \subset S^{n-1}$.
3. Using the data sets $\{\theta_i\}_{i=1}^l$ and $\{\eta_i\}_{i=1}^l$, derive a mapping $\hat{\varphi}(\cdot)$ which can approximate the homeomorphism $\varphi(\cdot)$ between the FPS boundary B_{P_N} and the $n-1$-sphere S^{n-1}.

3.2. **Construction of the mapping.**

LLE is a conceptually simple yet powerful manifold learning method. It maps the input data into a single global coordinate system of lower dimensionality. And the local geometry of input data is well preserved in the low dimensional space. Besides, LLE requires few parameters to be set and it avoids local minima inherent to many iterative techniques.

Actually, the FPS boundary B_{P_N} is an $n-1$ dimensional manifold. However, it is not easy to embed the sampled data θ_i, $i = 1, \cdots, l$ into $n-1$ dimensions by LLE, because the FPS boundary B_{P_N} is 'circular', i.e., has noncontractible loops [16]. Therefore, the data have to be embedded into n dimensions here.

In order to well sample the FPS boundary B_{P_N}, there should be sufficient uniformly sampled points θ_i. In [10], a method for computing the points on the FPS boundary was

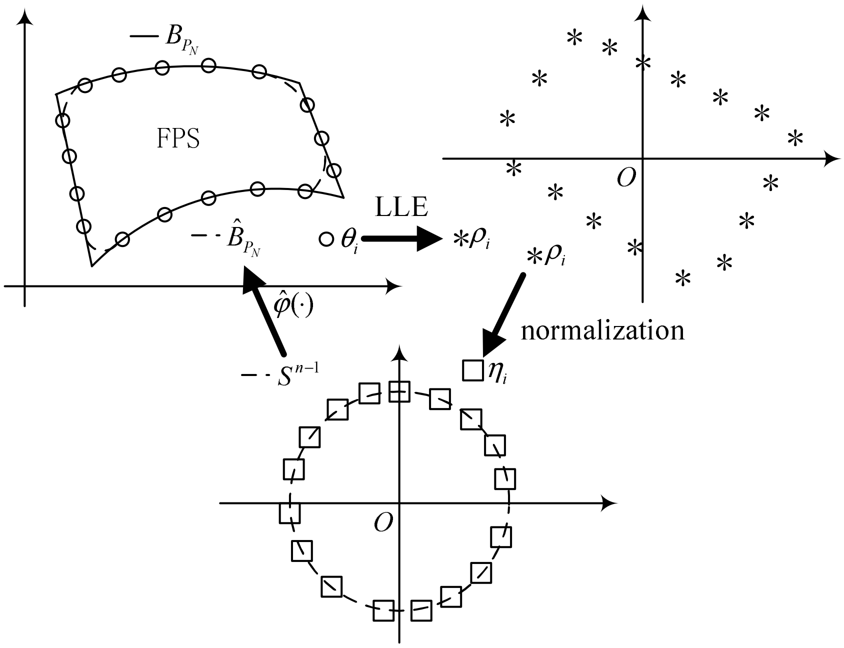


Figure 1. Construction of the mapping $\hat{\varphi}(\cdot)$

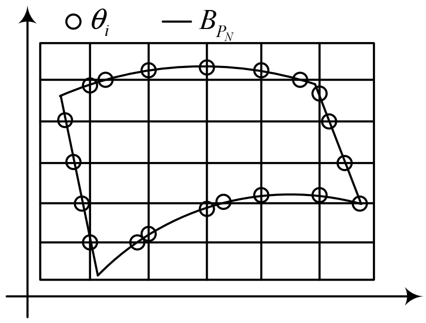


Figure 2. Sampling the FPS boundary B_{P_N}

proposed. However, this method suffers from the low efficiency. By this method, the number of the points to be computed is extremely large. This will increase the computational load of LLE. Furthermore, the obtained points are not uniformly distributed. This may deteriorate the performance of LLE. To overcome these difficulties, a simple but efficient strategy for sampling the FPS boundary B_{P_N} is proposed here. The strategy illustrated by Figure 2 consists of the following:

1. Obtain a box containing the FPS by prior knowledge or by the method given in [4].
2. Define a uniform grid covering the obtained box.
3. Find the intersections of the FPS boundary with the edges of the boxes of the grid by a no-derivative line search method.

Mapping the data set $\{\theta_i\}_{i=1}^{l}$ into a data set $\{\rho_i\}_{i=1}^{l}$ by LLE consists of the following:

1. Find K nearest neighbors of each point θ_i, $i = 1, \cdots, l$. The Euclidean distance is used as a similarity measure.

2. Compute the weights W_{ij} that best reconstruct each data point θ_i from its neighbors by minimizing the cost function

$$\sigma_1(W) = \sum_{i=1}^{l} \left(\left\| \theta_i - \sum_{j=1}^{l} W_{ij}\theta_j \right\|_2 \right)^2, \tag{9}$$

subject to constraints:

$$\sum\nolimits_{j=1}^{l} W_{ij} = 1 \tag{10}$$

and $W_{ij} = 0$ if θ_i and θ_j are not neighbors.

3. Compute the embedding ρ_i, $i = 1, \cdots, l$ best reconstructed by the weights W_{ij} by minimizing the cost function

$$\sigma_2(\mathrm{P}) = \sum_{i=1}^{l} \left(\left\| \rho_i - \sum_{j=1}^{l} W_{ij}\rho_j \right\|_2 \right)^2, \tag{11}$$

under constraints

$$\sum_{i=1}^{l} \rho_i = 0, \tag{12}$$

$$\frac{1}{l}\sum_{i=1}^{l} \rho_i {\rho_i}^{\mathrm{T}} = I, \tag{13}$$

where P is defined as $\mathrm{P} = [\rho_1, \rho_2, \cdots, \rho_l]$ and $I \in R^{n\times n}$ is a unit matrix. To find the matrix P under these constraints, a new matrix is constructed using the matrix W: $M = (I - W)^{\mathrm{T}}(I - W)$. LLE then computes the bottom $n + 1$ eigenvectors of M, associated with the $n+1$ smallest eigenvalues. The first eigenvector whose eigenvalue is zero is discarded. The remaining n eigenvectors yield the final embedding P.

The embedding of LLE is optimized to preserve the geometry of nearby inputs, but its shape is irregular. The constraints (12) and (13) require the embedded data to have zero mean and unit covariance. Hence, in order to obtain a data set with regular shape, the data set $\{\rho_i\}_{i=1}^{l}$ is mapped into a data set $\{\eta_i\}_{i=1}^{l} \subset S^{n-1}$: $\eta_i = \rho_i/\|\rho_i\|_2$, $i = 1, \cdots, l$.

After the first two steps, each vector $\theta_i \in B_{P_N}$ corresponds to a vector $\eta_i \in S^{n-1}$. However, in order to construct the mapping $\hat{\varphi}(\cdot)$, the computation of the vector θ corresponding to a new vector $\eta \in S^{n-1}$ is needed here. To cope with this problem, a non-parametric method based on the data sets $\{\theta_i\}_{i=1}^{l}$ and $\{\eta_i\}_{i=1}^{l}$ is used. This method computes the vector θ corresponding to a new vector $\eta \in S^{n-1}$ by the following:

1. Identify the K nearest neighbors of η among the vectors η_i, $i = 1, \cdots, l$.
2. Compute the linear weights w_i that best reconstruct η from its neighbors by minimizing the cost function $\sigma_3(w) = \left\| \eta - \sum_{i=1}^{l} w_i\eta_i \right\|_2$ subject to the constraints: $\sum_{i=1}^{l} w_i = 1$ and $w_i = 0$ if η_i is not one of the K nearest neighbors of η.
3. Compute the vector $\theta = \sum_{i=1}^{l} w_i\theta_i$, where the sum is over the vectors corresponding to the neighbors of η.

It is obvious that the construction procedure has only one free parameter to be determined: the number of neighbors K. Here, a method for determining this parameter is presented.

Denote by $\hat{\varphi}_K(\cdot)$ the obtained mapping after the implementation of the above construction procedure. Obtain a data set consisting of the vectors $\bar{\eta}_i$, $i = 1, \cdots, \bar{l}$, uniformly

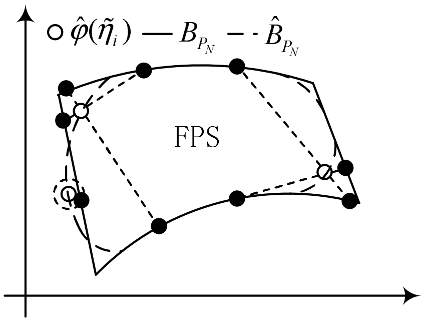


Figure 3. Searching for the new points on the boundary

sampled from the $n-1$-sphere S^{n-1}. Suppose that the number of the data is sufficient. Introduce a variable $\delta(K)$ describing the boundary approximation error:

$$\delta(K) = \max_i \left| \|Y - F(\bar{\theta}_i)\|_\infty^E - 1 \right|, \tag{14}$$

where

$$\bar{\theta}_i = \hat{\varphi}_K(\bar{\eta}_i), \quad i = 1, \cdots, \bar{l}. \tag{15}$$

The smaller $\delta(K)$, the better the boundary approximation. Hence, the optimal number of neighbors K^* should be $\arg\min_K \delta(K)$. Finally, the constructed mapping $\hat{\varphi}(\cdot)$ is set to be $\hat{\varphi}_{K^*}(\cdot)$.

3.3. **Improvement of the boundary approximation.** It is not difficult to find that the mapping $\hat{\varphi}(\cdot)$ is only an approximation of the homeomorphism $\varphi(\cdot)$ between the FPS boundary B_{P_N} and the $n-1$-sphere S^{n-1}. However, once the mapping $\hat{\varphi}(\cdot)$ is constructed, it can be used to map the $n-1$-sphere S^{n-1} into a set $\hat{B}_{P_N}$ approximating the FPS boundary B_{P_N}.

Suppose that there are sufficient points $\tilde{\eta}_i$, $i = 1, \cdots, \tilde{l}$ uniformly sampled from the $n-1$-sphere S^{n-1}. By the mapping $\hat{\varphi}(\cdot)$, a set of the points $\hat{\varphi}(\tilde{\eta}_i)$, $i = 1, \cdots, \tilde{l}$ on or near the FPS boundary B_{P_N} can be obtained. This set can thus be used to approximately characterize the FPS boundary B_{P_N}. Here, a simple strategy is presented to further better the boundary approximation. The strategy consists of the following:

1. Find the points $\hat{\varphi}(\tilde{\eta}_i)$ whose approximation errors $\left| \|Y - F(\hat{\varphi}(\tilde{\eta}_i))\|_\infty^E - 1 \right|$ are bigger than a prespecified threshold γ.
2. For each point found in step 1, if it is outside the FPS, then within a small neighborhood of it, find a point on the FPS boundary, else, from it in several randomly chosen directions, search for the points on the FPS boundary and then search in the opposite directions as well. (This step is illustrated by Figure 3.)
3. Discard the points found in step 1. Use a set of all the remaining points (including the ones found in step 2) to approximately characterize the FPS boundary B_{P_N}.

4. Examples.

4.1. **Sine function.** Consider the regression model

$$y_k = \sin(p_1 x_k) + p_2 + e_k, \tag{16}$$

where p_1 and p_2 are the elements of the parameter vector. Two measurements are obtained: for $x_1 = 1$, $y_1 = 1$; for $x_2 = 3$, $y_2 = 0.5$. Suppose that $|e_k| \le 0.5$, $k = 1, 2$. From the error bounds, it is obtained that $E = [0.5\ \ 0.5]^{\mathrm{T}}$. Besides, a prior knowledge about the parameters is assumed to be available: $p_1 \in [2.8, 4.2]$ and $p_2 \in [0, 2.2]$.

A uniform grid covering $[2.8, 4.2] \times [0, 2.2]$ is defined. A set of 64 points is obtained by computing the intersections of the FPS boundary with the edges of boxes of the grid which are parallel to p_2-axis. Figure 4 illustrates the sampled data. A set of 400 points uniformly sampled from 1-sphere S^1 ($\bar{l} = 400$) is used in Equations (14) and (15) for obtaining the best number of neighbors K. Since $\arg\min_K \delta(K) = 5$, set K to be 5. Figure 5 shows δ versus K. Figure 6(a) shows the embedding given by LLE. Figure 6(b) illustrates the normalized data. From the figure, it can be observed that not only the sampled points can be mapped into the points on the 1-sphere, but also the local geometry is well preserved. Figures 7(a) and 7(b) show the approximations of the FPS boundary given by the proposed method and the support vector machine (SVM) method [12], respectively. For the latter, a set of 117 points uniformly sampled in $[2.8, 4.2] \times [0, 2.2]$ is used, and an LS-SVM with the RBF kernel whose width is equal to $\sqrt{0.2}$ is utilized to obtain the best result. It can be observed that the proposed method can give a better approximation than the SVM method. Besides, it is apparent that a single convex set containing the FPS is very pessimistic.

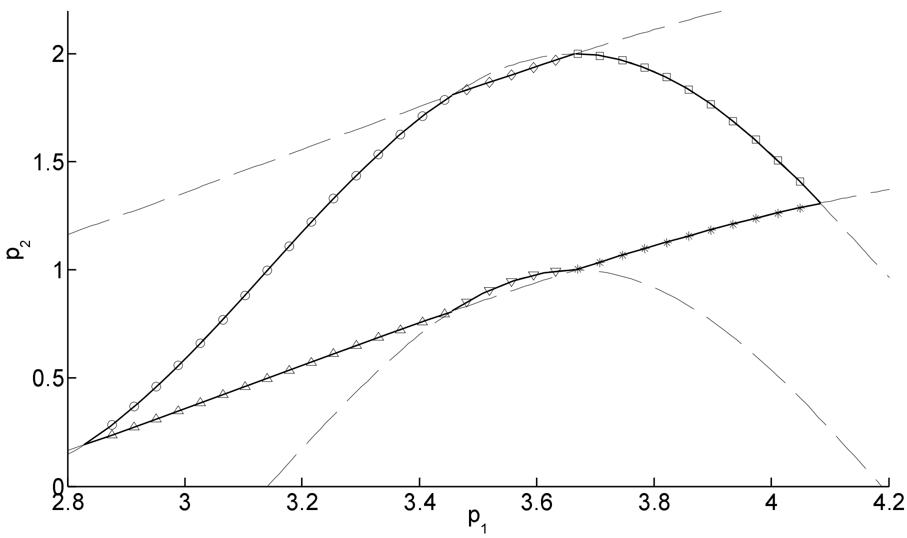


FIGURE 4. Points sampled from the FPS boundary

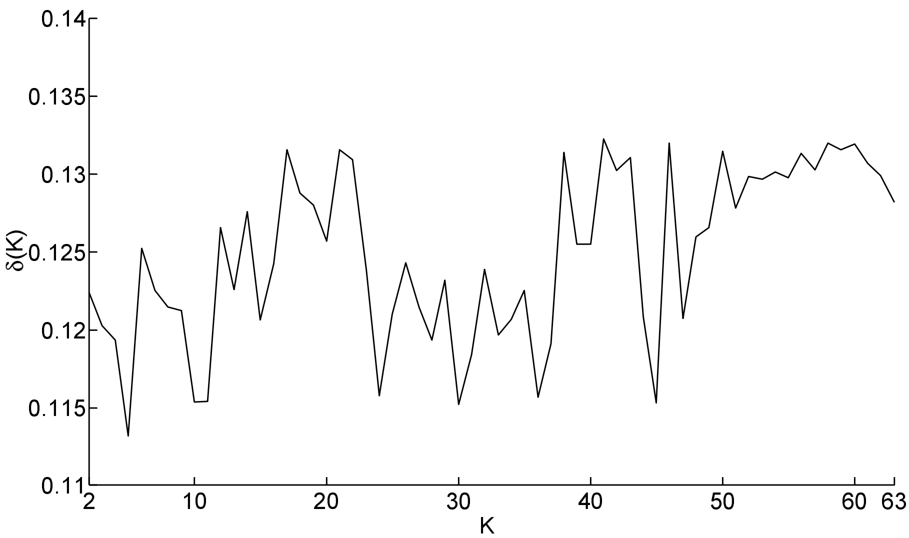


FIGURE 5. Error δ versus the number of neighbors K

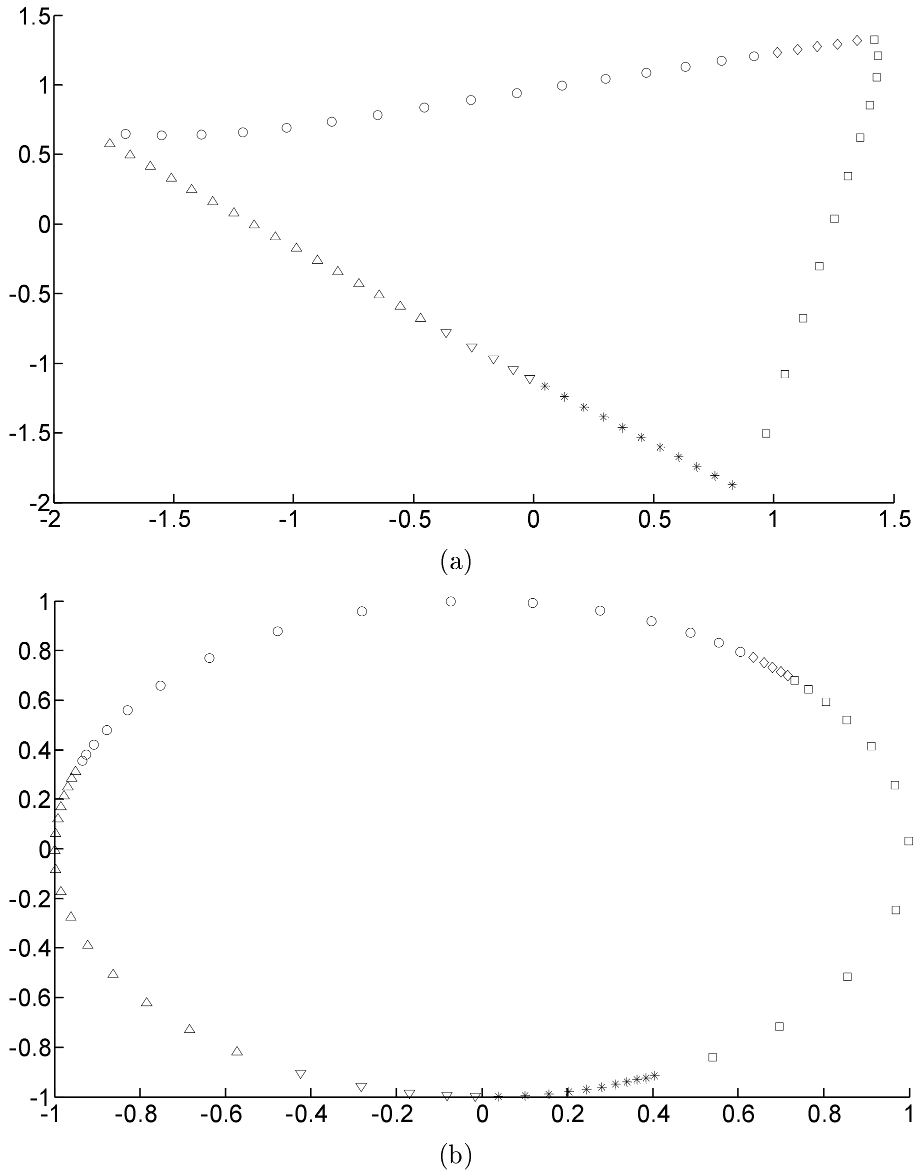


FIGURE 6. Results of LLE and data normalization: (a) the embedding given by LLE; (b) the normalized data

4.2. **Exponential function.** Consider the regression model

$$y_k = p_2 \exp(-p_1 x_k) + e_k, \tag{17}$$

where p_1 and p_2 are the elements of the parameter vector. Two measurements are obtained: for $x_1 = 1$, $y_1 = 0.15$; for $x_2 = 5$, $y_2 = 0.03$. Suppose that $|e_1| \leq 0.1$ and $|e_2| \leq 0.01$. From the error bounds, it is obtained that $E = [0.1 \ \ 0.01]^{\mathrm{T}}$.

A box $[0, 0.65] \times [0, 0.5]$ containing the FPS is obtained by the method in [4]. A uniform grid covering this box is defined. A set of 68 points is obtained by computing the intersections of the FPS boundary with the edges of boxes of the grid which are parallel to p_2-axis. Figure 8 illustrates the sampled data. A set of 400 points ($\bar{l} = 400$) is used to obtain the best number of neighbors K. Since $\arg\min_K \delta(K) = 4$, set K to be 4. Figures 9(a) and 9(b) show the approximations of the FPS boundary. For the SVM method, a set of 441 points uniformly sampled in $[0, 0.65] \times [0, 0.5]$ is used, and an LS-SVM with the width of RBF kernel equal to $\sqrt{0.01}$ is utilized to obtain the best result. It can be

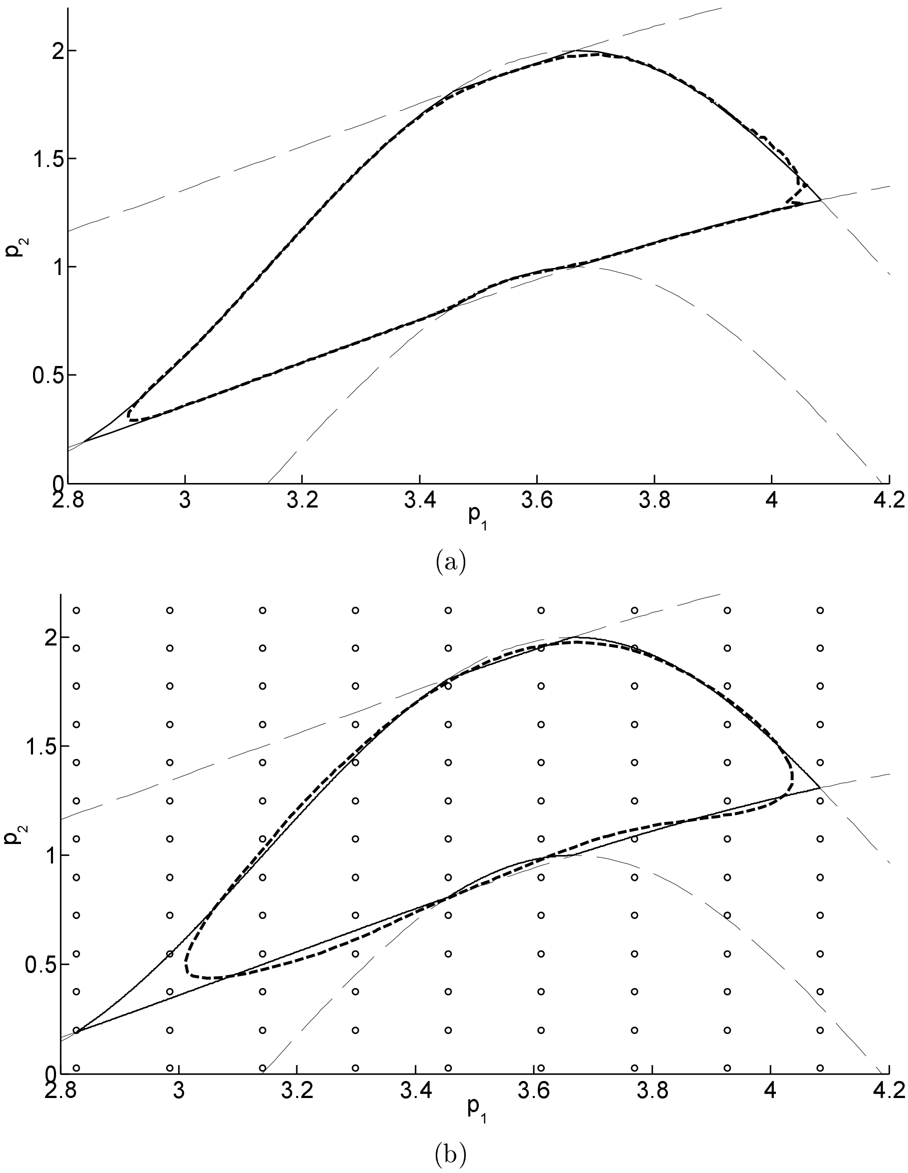


FIGURE 7. Approximations of the FPS boundary given by (a) the proposed method and (b) the SVM method. Solid line: exact boundary. Dashed line: approximation of the boundary.

observed that the proposed method has better approximation accuracy. And a single convex set containing the FPS is apparently pessimistic.

Figure 10 shows the improvement of the boundary approximation. A set of 80 points uniformly sampled from 1-sphere S^1 ($\tilde{l} = 80$) is obtained. The set of the images of these points under the mapping $\hat{\varphi}(\cdot)$ is used to approximately characterize the FPS boundary. Nine points whose approximation errors are bigger than a threshold 0.015 are found. Seven of the nine points are in the FPS while the rest are outside. For each of the seven points, three pairs of opposite search directions are randomly chosen. A set of 115 points on or near the FPS boundary is obtained as the final result. For comparison, Figure 11 shows the result of the method in [10] when the number of the points on the FPS boundary to be found is set to be 120. For this method, the origin of searching is initially set to be $[0.58\ 0.4]^{\mathrm{T}}$, and three pairs of opposite search directions are randomly chosen

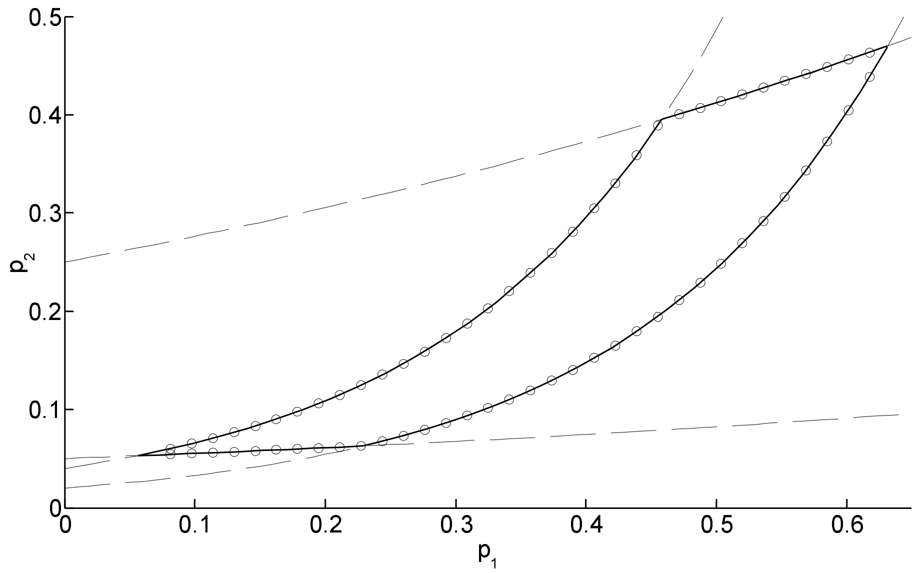


FIGURE 8. Points sampled from the FPS boundary

for each origin. It can be observed that the proposed method has higher efficiency than Lahanier's method.

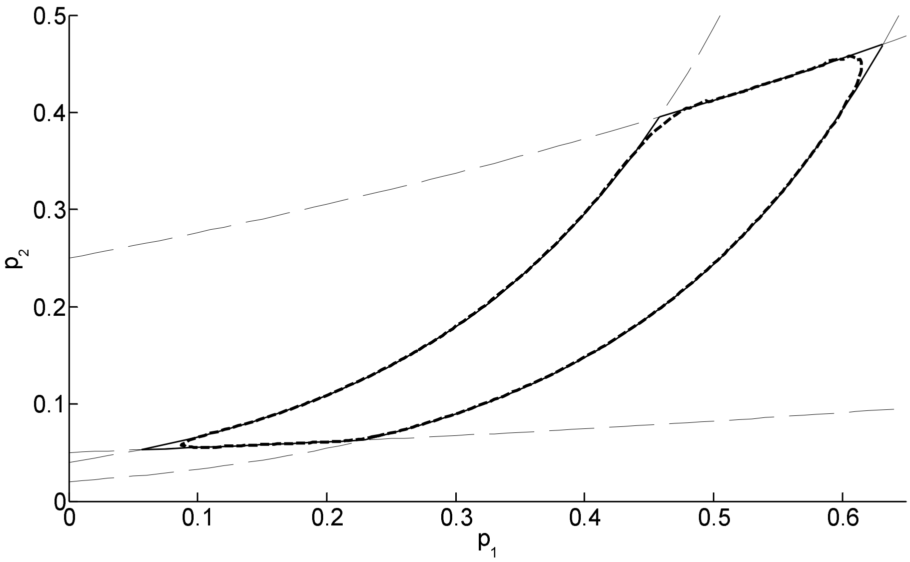
4.3. **Pharmacokinetic model.** Consider the classical one-compartment open linear model with first-order absorption described by

$$y_k = \frac{\text{Dose} \cdot K_a}{V(K_a - K_e)} [\exp(-K_e x_k) - \exp(-K_a x_k)] + e_k, \tag{18}$$

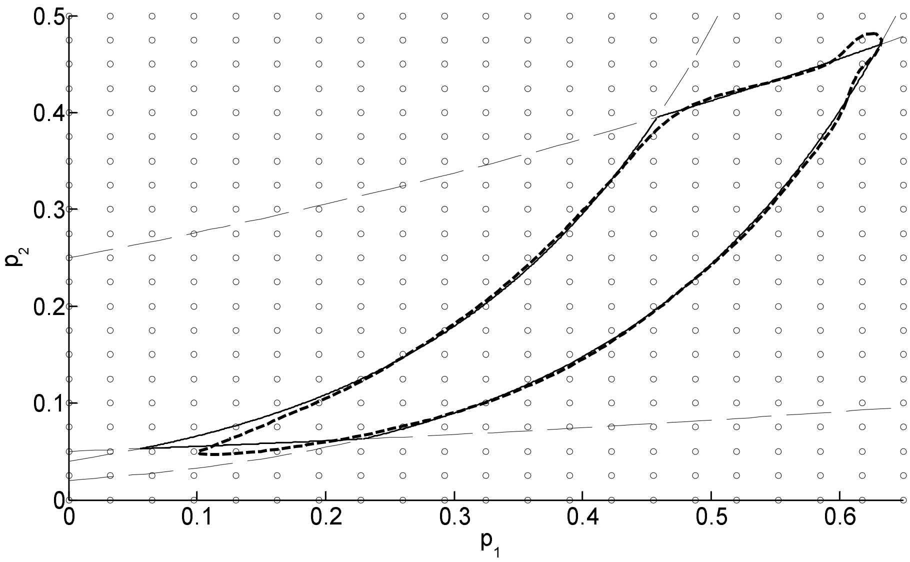
where y_k is the observed drug concentration at time x_k, K_a is the absorption rate constant ($K_a = 1.1\text{hr}^{-1}$), K_e is the elimination rate constant ($K_e = 0.2\text{hr}^{-1}$), and V is the volume of distribution ($V = 50$L). A data set is obtained by measuring the drug concentration in 12 plasma samples at instants 0.25, 0.5, 1, 1.5, 2, 4, 6, 8, 10, 12, 18, and 24hr using an oral dose of 1000mg of a drug. The error e_k is truncated normally distributed with mean 0mg/L and standard deviation 0.2/3mg/L. The error bound is 0.2mg/L. From the error bound, it is obtained that $E_i = 0.2$, $i = 1, 2, \cdots, 12$. K_a and K_e are the parameters to be estimated.

A box $[1.05, 1.14] \times [0.194, 0.208]$ containing the FPS is obtained by the method in [4]. A uniform grid covering this box is defined. A set of 78 points is obtained by the sampling strategy. Figure 12 illustrates the sampled data. A set of 400 points ($\bar{l} = 400$) is used to obtain the best number of neighbors K. Since $\arg\min_K \delta(K) = 27$, set K to be 27. Figures 13(a) and 13(b) show the approximations of the FPS boundary. For the SVM method, a set of 441 points uniformly sampled in $[1.05, 1.14] \times [0.194, 0.208]$ is used, and an LS-SVM with the width of RBF kernel equal to $\sqrt{1.4 \times 10^{-5}}$ is utilized to obtain the best result. It can be seen that the proposed method has better approximation accuracy.

Figure 14 shows the improvement of the boundary approximation. A set of 80 points ($\tilde{l} = 80$) is used to approximately characterize the FPS boundary. Thirteen points whose approximation errors are bigger than a threshold 0.0005 are found. Eleven of the thirteen points are in the FPS while the rest are outside. For each of the eleven points, two pairs of opposite search directions are randomly chosen. A set of 113 points on or near the FPS boundary is obtained as the final result. Figure 15 shows the result of Lahanier's method when the number of the points on the FPS boundary to be found is set to be 120. For this method, the origin of searching is initially set to be $[1.1\ 0.2]^{\mathrm{T}}$, and two pairs of opposite



(a)



(b)

FIGURE 9. Approximations of the FPS boundary given by (a) the proposed method and (b) the SVM method. Solid line: exact boundary. Dashed line: approximation of the boundary.

search directions are randomly chosen for each origin. It can be seen that the proposed method has higher efficiency than Lahanier's method.

5. **Conclusions.** This paper proposes a novel set membership identification method for models nonlinear in their parameters. This method can deal with complex models for which it is difficult to construct the minimal inclusion function. This paper seeks a mapping which can approximate the homeomorphism between the FPS boundary and the $n-1$-sphere (n is the number of parameters). A method for constructing this mapping is presented. Once this mapping is established, it can be used to map the $n-1$-sphere into an approximation of the FPS boundary. Besides, a strategy for improving the boundary approximation is proposed. This strategy discards the unsatisfied points and from them

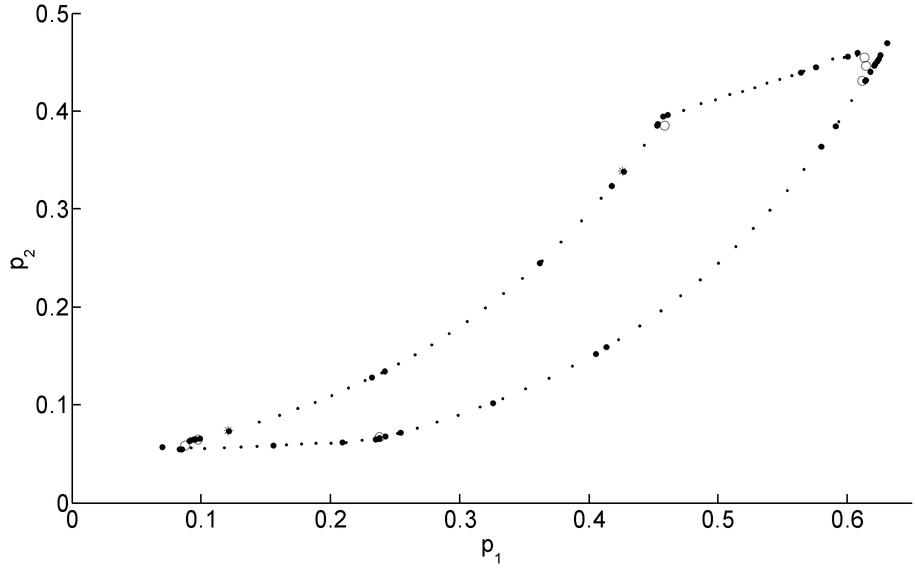


FIGURE 10. Improvement of the boundary approximation. Small points: points whose approximation errors are smaller than the threshold. Circles: among the points whose approximation errors are bigger than the threshold, the ones which are in the FPS. Asterisks: among the points whose approximation errors are bigger than the threshold, the ones which are outside the FPS. Big points: points obtained by searching.

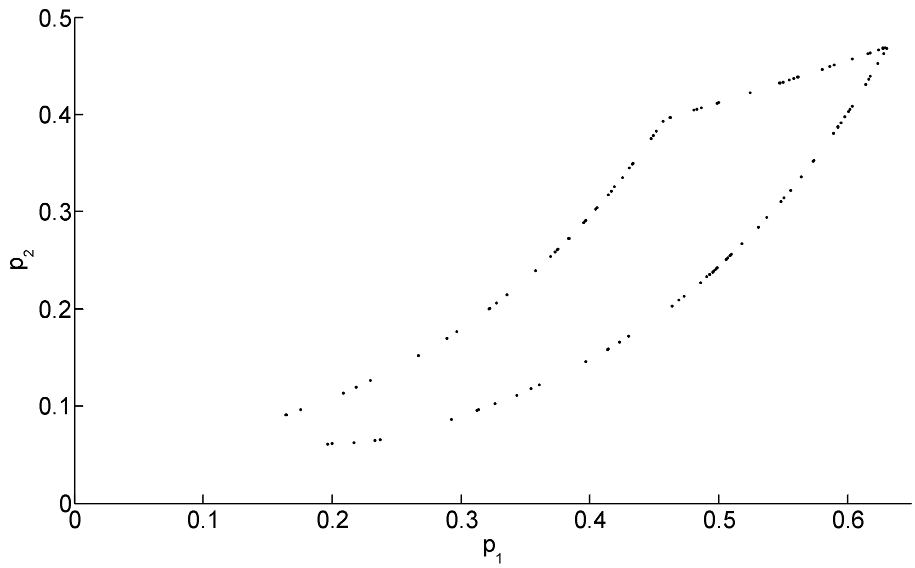


FIGURE 11. Result of Lahanier's method

searches for new points on the FPS boundary. Three examples demonstrate the performance of the proposed method. It can give a less pessimistic result than the methods using a single convex set to characterize the FPS. Moreover, the proposed method can give a better approximation than the SVM method. Compared with Lahanier's method, the proposed method needs to compute fewer points on the FPS boundary, i.e., it has higher efficiency. This paper assumes the FPS to be connected. The future work will concentrate on the extension of the proposed method to the description of nonconnected FPS.

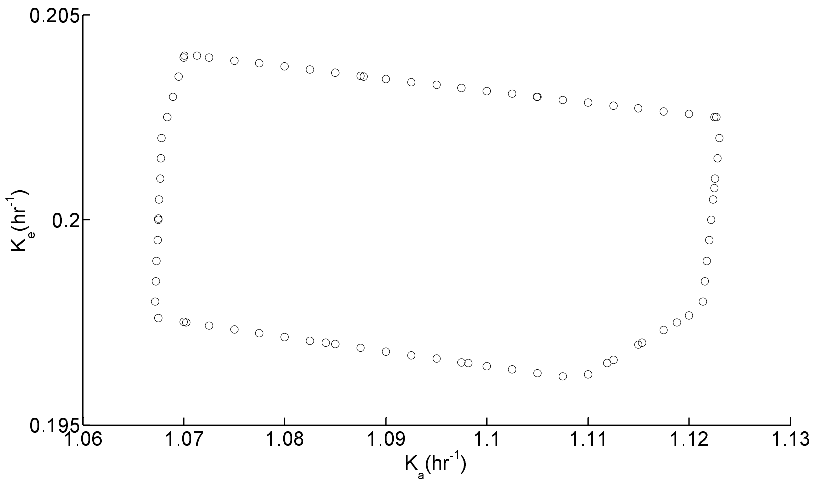
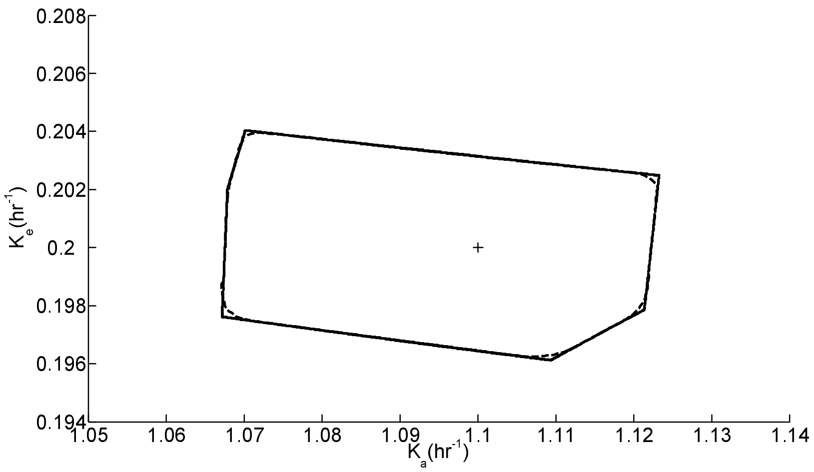
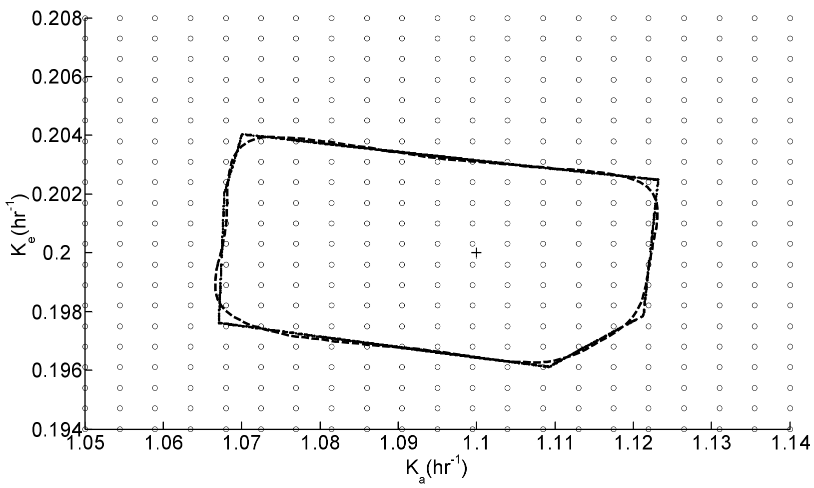


FIGURE 12. Points sampled from the FPS boundary



(a)



(b)

FIGURE 13. Approximations of the FPS boundary given by (a) the proposed method and (b) the SVM method. Dotted line: exact boundary. Dashed line: approximation of the boundary. Plus sign: True value of the parameter vector.

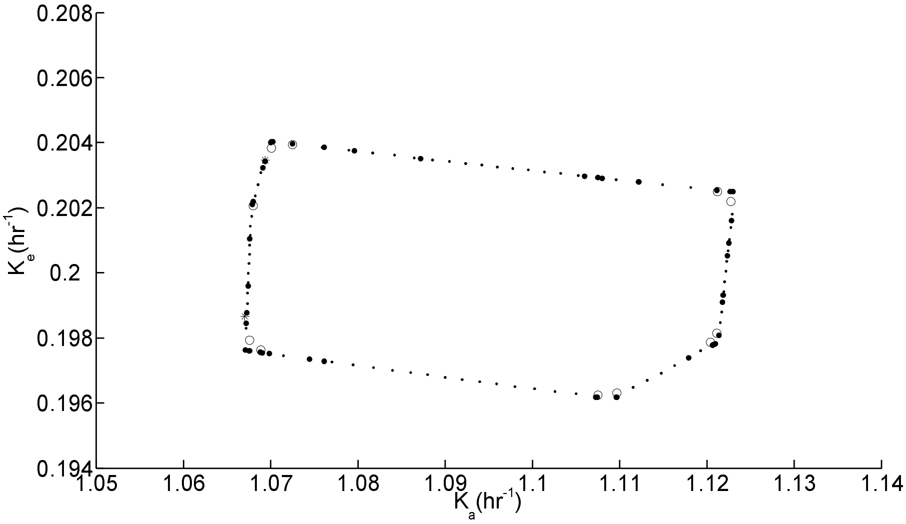


Figure 14. Improvement of the boundary approximation. Small points: points whose approximation errors are smaller than the threshold. Circles: among the points whose approximation errors are bigger than the threshold, the ones which are in the FPS. Asterisks: among the points whose approximation errors are bigger than the threshold, the ones which are outside the FPS. Big points: points obtained by searching.

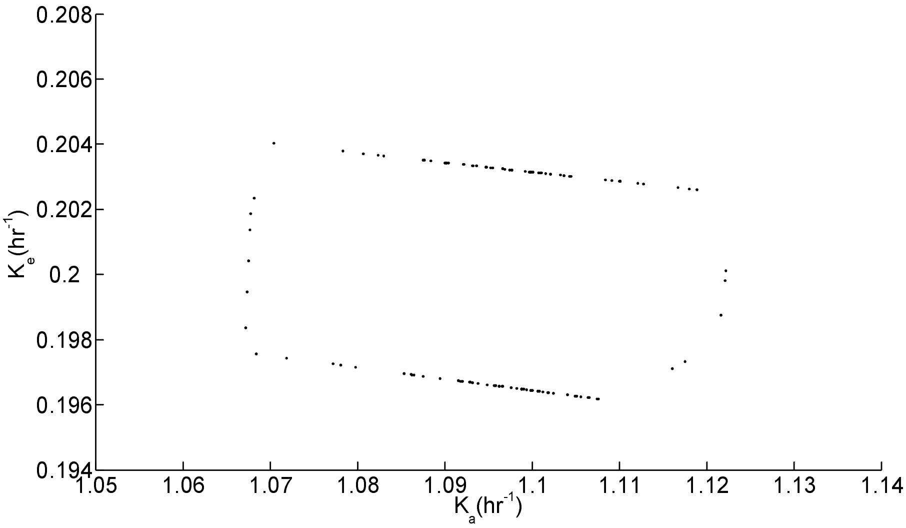


Figure 15. Result of Lahanier's method

Acknowledgment. The authors wish to thank the anonymous reviewers for their constructive comments and suggestions. This work is supported by the Beijing Municipal Natural Science Fund (Grant No. 4144067).

REFERENCES

[1] E. Walter and H. Piet-Lahanier, Estimation of parameter bounds from bounded-error data: A survey, *Mathematics and Computers in Simulation*, vol.32, no.5-6, pp.449-468, 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] E. Walter, J. Norton, H. Piet-Lahanier and M. Milanese, *Bounding Approaches to System Identification*, Plenum Press, New York, 1996.

[4] M. Milanese and A. Vicino, Estimation theory for nonlinear models and set membership uncertainty, *Automatica*, vol.27, no.2, pp.403-408, 1991.
[5] E. Scholte and M. E. Campbell, A nonlinear set-membership filter for on-line applications, *International Journal of Robust and Nonlinear Control*, vol.13, no.15, pp.1337-1358, 2003.
[6] W. Chai and X. F. Sun, Nonlinear set membership filtering using ellipsoids and its application in fault diagnosis, *Acta Aeronautica et Astronautica Sinica*, vol.28, no.4, pp.948-952, 2007 (in Chinese).
[7] J. M. Bravo, T. Alamo, M. J. Redondo and E. F. Camacho, An algorithm for bounded-error identification of nonlinear systems based on DC functions, *Automatica*, vol.44, no.2, pp.437-444, 2008.
[8] L. Jaulin and E. Walter, Set inversion via interval analysis for nonlinear bounded-error estimation, *Automatica*, vol.29, no.4, pp.1053-1064, 1993.
[9] T. Raïssi, N. Ramdani and Y. Candau, Set membership state and parameter estimation for systems described by nonlinear differential equations, *Automatica*, vol.40, no.10, pp.1771-1777, 2004.
[10] H. Lahanier, E. Walter and R. Gomeni, OMNE: A new robust membership-set estimator for the parameters of nonlinear models, *Journal of Pharmacokinetics and Biopharmaceutics*, vol.15, no.2, pp.203-219, 1987.
[11] K. Keesman, Membership-set estimation using random scanning and principal component analysis, *Mathematics and Computers in Simulation*, vol.32, no.5-6, pp.535-543, 1990.
[12] K. J. Keesman and R. Stappers, Nonlinear set-membership estimation: A support vector machine approach, *Journal of Inverse and Ill-Posed Problems*, vol.12, no.1, pp.27-41, 2004.
[13] W. Chai and X. F. Sun, Set membership estimation by weighted least squares support vector machines, *Electric Machines and Control*, vol.13, no.3, pp.431-435, 2009 (in Chinese).
[14] S. T. Roweis and L. K. Saul, Nonlinear dimensionality reduction by locally linear embedding, *Science*, vol.290, no.5500, pp.2323-2326, 2000.
[15] L. K. Saul and S. T. Roweis, Think globally, fit locally: Unsupervised learning of low dimensional manifolds, *Journal of Machine Learning Research*, vol.4, no.2, pp.119-155, 2003.
[16] J. A. Lee and M. Verleysen, Nonlinear dimensionality reduction of data manifolds with essential loops, *Neurocomputing*, vol.67, pp.29-53, 2005.