

Online identification of nonlinear system using reduced kernel principal component analysis

Okba Taouali · Ilyes Elaïssi · Hassani Messaoud

Received: 9 November 2009 / Accepted: 4 October 2010 / Published online: 23 October 2010
© Springer-Verlag London Limited 2010

Abstract The Principal Component Analysis (PCA) is a powerful technique for extracting structure from possibly high-dimensional data sets. It is readily performed by solving an eigenvalue problem, or by using iterative algorithms that estimate principal components. This paper proposes a new method for online identification of a nonlinear system modelled on Reproducing Kernel Hilbert Space (RKHS). Therefore, the PCA technique is tuned twice, first we exploit the Kernel PCA (KPCA) which is a nonlinear extension of the PCA to RKHS as it transforms the input data by a nonlinear mapping into a high-dimensional feature space to which the PCA is performed. Second, we use the Reduced Kernel Principal Component Analysis (RKPCA) to update the principal components that represent the observations selected by the KPCA method.

Keywords RKHS · SLT · Kernel method · RKPCA · Online RKPCA

1 Introduction

Since the introduction of Support Vector Machines (SVM) [19], many learning algorithms have been transferred to a kernel representation [2, 7]. The important benefit lies on

the fact that nonlinearities can be allowed, while avoiding to solve a nonlinear optimization problem. The transfer is implicitly accomplished by means of a nonlinear map in a Reproducing Kernel Hilbert Space F_k (RKHS) [9].

Kernel methods have been successfully applied to a large class of problems, such as identification of nonlinear system [2, 3, 16, 17], diagnostic system [24], time series prediction [13], face recognition [26], biological data processing for medical diagnosis [23]... The attractiveness of such algorithms stands from their elegant treatment of data issued from nonlinear processes. However, these techniques suffer from computational complexity as the amount of computer memory and the training time increases rapidly with the number of observations. It is clear that for large datasets (as for example in image processing, computer vision or object recognition), the kernel method with its powerful advantage of dealing with nonlinearities is computationally limited. For large datasets, an eigen decomposition of Gram matrix can simply become too time-consuming to extract the principal components and therefore the system parameter identification becomes a tough task. To overcome this burden, recently a theoretical foundation for online learning algorithm with kernel method in reproducing kernel Hilbert spaces was proposed [4, 5, 14, 15, 18, 22, 25]. Also online kernel algorithm is more useful when the system to be identified is time-varying, because these algorithms can automatically track changes of system model with time-varying and time lagging characteristic.

In this paper, we propose a new method for online identification of a nonlinear system parameters modeled on Reproducing Kernel Hilbert Space (RKHS). This method uses the Reduced Kernel Principal Component Analysis (RKPCA) that selects the observation data to approach the Principal Components Analysis kept by the Kernel

O. Taouali (✉) · I. Elaïssi · H. Messaoud
Unité de Recherche d'Automatique,
Traitement de Signal et Image (ATSI),
5000 Monastir, Tunisia
e-mail: taoualiokba@yahoo.fr

I. Elaïssi
e-mail: ilyes.elaissi@yahoo.fr

H. Messaoud
e-mail: hassani.messaoud@enim.rnu.tn

Principal Component Analysis method KPCA [2]. The selected observations are used to build an RKHS model with a reduced parameter number. The proposed online identification method updates the list of the retained principal components and then the RKHS model by evaluating the error between the output model and the process one. The proposed technique may be very helpful to design an adaptive control strategy of nonlinear systems.

The paper is organized as follows. In Sect. 2, we remind the Reproducing Kernel Hilbert Space (RKHS). Section 3 is devoted to the modeling in RKHS. The Reduced Kernel Principal Component Analysis RKPCA method is presented in Sect. 4. In Sect. 5, we propose the new online RKPCA method. The proposed algorithm has been tested to identify the Wiener-Hammerstein benchmark [21] and a chemical reactor [8].

2 Reproducing kernel Hilbert space

Let $E \subset \mathbb{R}^d$ be an input space and $L^2(E)$ the Hilbert space of square integrable functions defined on E . Let $k : E \times E \rightarrow \mathbb{R}$ be a continuous positive definite kernel. It is proved [6, 9] that it exists as a sequence of an orthonormal eigen functions $(\psi_1, \psi_2, \dots, \psi_l)$ in $L^2(E)$ and a sequence of corresponding real positive eigenvalues $(\sigma_1, \sigma_2, \dots, \sigma_l)$ (where l can be infinite) so that

$$k(x, t) = \sum_{j=1}^l \sigma_j \psi_j(x) \psi_j(t); \quad x, t \in E. \quad (1)$$

Let $F_k \subset L^2(E)$ be a Hilbert space associated to the kernel k and defined by:

$$F_k = \left\{ f \in L^2(E) / f = \sum_{i=1}^l w_i \varphi_i \text{ and } \sum_{j=1}^l \frac{w_j^2}{\sigma_j} < +\infty \right\} \quad (2)$$

where $\varphi_i = \sqrt{\sigma_i} \psi_i$ $i = 1, \dots, l$. The scalar product in the space F_k is given by:

$$\langle f, g \rangle_{F_k} = \left\langle \sum_{i=1}^l w_i \varphi_i, \sum_{j=1}^l z_j \varphi_j \right\rangle_{F_k} = \sum_{i=1}^l w_i z_i \quad (3)$$

The kernel k is said to be a reproducing kernel of the Hilbert space F_k if and only if the following conditions are satisfied.

$$\begin{cases} \forall x \in E, & k(x, \cdot) \in F_k \\ \forall x \in E & \text{and } \forall f \in F_k, \langle f(\cdot), k(x, \cdot) \rangle_{F_k} = f(x) \end{cases} \quad (4)$$

where $k(x, \cdot)$ means $k(x, x')$ $\forall x' \in E$. F_k is called reproducing kernel Hilbert space (RKHS) with kernel k and dimension l . Moreover, for any RKHS, there exists only one positive definite kernel and vice versa [10].

Among the possible reproducing kernels, we mention the Radial Basis function (RBF) defined as:

$$k(x, t) = \exp\left(-\|x - t\|^2 / 2\mu^2\right); \quad \forall x, t \in E \quad (5)$$

with μ a fixed parameter.

3 RKHS models

Consider a set of observations $\{x^{(i)}, y^{(i)}\}_{i=1, \dots, M}$ with $x^{(i)} \in \mathbb{R}^n, y^{(i)} \in \mathbb{R}$ are respectively the system input and output. According to the statistical learning theory (SLT) [19, 20], the identification problem in the RKHS F_k can be formulated as a minimization of the regularized empirical risk. Thus, it consists in finding the function $f^* \in F_k$ such that

$$f^* = \sum_{j=1}^l w_j^* \varphi_j = \min_{f \in F_k} \frac{1}{M} \sum_{i=1}^M \left(y^{(i)} - f(x^{(i)}) \right)^2 + \lambda \|f\|_{F_k}^2 \quad (6)$$

where M is the measurement number and λ is a regularization parameter chosen in order to ensure a generalization ability to the solution f^* . According to the representer theorem [9], the solution f^* of the optimization problem (6) is a linear combination of the kernel k applied to the M measurements $x^{(i)}$, $i = 1, \dots, M$, as:

$$f^*(x) = \sum_{i=1}^M \alpha_i^* k(x, x^{(i)}), x. \quad (7)$$

To solve the optimization problem (6), we can use some kernel methods such that Support Vector Machine (SVM) [11], Least Square Support Vector Machine (LSSVM) [7], Regularization Network (RN) [3], Kernel Partial Least Square (KPLS) [12], In [2], the Kernel Principal Component Analysis KPCA were proposed. This method reconsiders the regularization idea by finding the solution to the identification problem in some subspace F_{kPCA} spanned by the so-called principal component and yields to a RKHS model with M parameters.

In the next section, we present the Reduced KPCA method in which we approximate the retained principal components given by the KPCA with a set of vectors of input observations. This approximation is performed by a set of particular training observations and allows the construction of a RKHS model with much less parameters.

4 RKPCA method

Let a nonlinear system with an input $u \in \mathbb{R}$ and an output $y \in \mathbb{R}$ from which we extract a set of observations be $\{u^{(i)}, y^{(i)}\}_{i=1, \dots, M}$. Let F_k be an RKHS space with kernel k .

To build the input vector $x^{(i)}$ of the RKHS model, we use the NARX (Nonlinear auto regressive with eXogeneous input) structure as:

$$x^{(i)} = \left\{ u^{(i)}, \dots, u^{(i-m_u)}, y^{(i-1)}, \dots, y^{(i-m_y)} \right\}^T; \quad m_u, m_y \in \mathbb{N} \tag{8}$$

The set of observations becomes $D = \{x^{(i)}, y^{(i)}\}_{i=1, \dots, M}$ where $x^{(i)} \in \mathbb{R}^{m_u+m_y+1}$ and $y^{(i)} \in \mathbb{R}$ and the RKHS model of this system based on (7) can be written as:

$$\tilde{y}^{(j)} = \sum_{i=1}^M a_i k(x^{(i)}, x^{(j)}) \tag{9}$$

Let the application Φ :

$$\Phi : E \rightarrow \mathbb{R}^l$$

$$x \mapsto \Phi(x) = \begin{pmatrix} \varphi_1(x) \\ \vdots \\ \varphi_l(x) \end{pmatrix} \tag{10}$$

where φ_i are given in (2).

The Gram matrix K associated with the kernel k is an M -dimensional square matrix so that

$$K_{i,j} = k(x^{(i)}, x^{(j)}) \quad \text{for } i, j = 1, \dots, M \tag{11}$$

The kernel trick [10] is so that

$$\langle \Phi(x), \Phi(x') \rangle = k(x, x') \quad x, x' \in E \tag{12}$$

We assume that the transformed data $\{\Phi(x^{(i)})\}_{i=1, \dots, M} \in \mathbb{R}^l$ are centered [2]. The empirical covariance matrix of the transformed data is symmetrical and l -dimensional. It is written as following:

$$C_\phi = \frac{1}{M} \sum_{i=1}^M \Phi(x^{(i)}) \Phi(x^{(i)})^T, \quad C_\phi \in \mathbb{R}^{l \times l} \tag{13}$$

Let l' be the number of the eigenvectors $\{V_j\}_{j=1, \dots, l'}$ of the C_ϕ matrix that corresponds to the nonzeros positive eigenvalues $\{\lambda_j\}_{j=1, \dots, l'}$. It is proved in [2] that the number l' is less or equal to M .

Due to the large size l of C_ϕ , the calculus of $\{V_j\}_{j=1, \dots, l'}$ can be difficult. The KPCA method shows that these $\{V_j\}_{j=1, \dots, l'}$ are related to the eigenvectors $\{\beta_j\}_{j=1, \dots, l'}$ of the gram matrix K according to [1]:

$$V_j = \sum_{i=1}^M \beta_{j,i} \Phi(x^{(i)}), \quad j = 1, \dots, l' \tag{14}$$

where $(\beta_{j,i})_{j=1, \dots, l'}$ are the components of $\{\beta_j\}_{j=1, \dots, l'}$ associated to their nonzero eigenvalues $\mu_1 > \dots > \mu_{l'}$.

The principle of the KPCA method consists in organizing the eigenvectors $\{\beta_j\}_{j=1, \dots, l'}$ in the decreasing order

of their corresponding eigenvalues $\{\mu_j\}_{j=1, \dots, l'}$. The principal components are the p first vectors $\{V_j\}_{j=1, \dots, p}$ associated with the highest eigenvalues and are often sufficient to describe the structure of the data [1, 2]. The number p satisfies the Inertia Percentage criterion IPC given by:

$$p^* = \arg(\text{IPC} \geq 99) \tag{15}$$

where

$$\text{IPC} = \frac{\sum_{i=1}^p \mu_i}{\sum_{i=1}^M \mu_i} \times 100 \tag{16}$$

The RKHS model provided by the KPCA method is [1].

$$\tilde{y}^{(\text{new})} = \sum_{q=1}^p w_q \sum_{i=1}^M \beta_{q,i} k(x^i, x^{(\text{new})}) \tag{17}$$

Since the principal components are a linear combination of the transformed input data $\{\Phi(x^i)\}_{i=1, \dots, M}$ [3], the Reduced KPCA approaches each vector $\{V_j\}_{j=1, \dots, p}$ by a transformed input data $\Phi(x_b^{(j)}) \in \{\Phi(x^i)\}_{i=1, \dots, M}$ having a high projection value in the direction of V_j [1].

The projection of the $\Phi(x^{(i)})$ on the V_j called $\tilde{\Phi}(x^{(i)})_j \in \mathbb{R}$ and can be written as:

$$\tilde{\Phi}(x^{(i)})_j = \langle V_j, \Phi(x^{(i)}) \rangle, \quad j = 1, \dots, p \tag{18}$$

According to (14) and (12), the relation (18) is written:

$$\tilde{\Phi}(x^{(i)})_j = \sum_{m=1}^M \beta_{j,m} k(x^{(m)}, x^{(i)}), \quad j = 1, \dots, p \tag{19}$$

To select the vectors $\{\Phi(x_b^{(i)})\}$, we project all the $\{\Phi(x^{(i)})\}_{i=1, \dots, M}$ vectors on each principal component $\{V_j\}_{j=1, \dots, p}$ and we retained $x_b^{(j)} \in \{x^{(i)}\}_{i=1, \dots, M}$ that satisfies

$$\begin{cases} \Phi(x_b^{(j)})_j = \text{Max}_{i=1, \dots, M} \tilde{\Phi}(x^{(i)})_j \\ \text{and} \\ \Phi(x_b^{(j)})_{i \neq j} < \zeta \end{cases} \tag{20}$$

where ζ is a given threshold.

Once the $\{x_b^{(j)}\}_{j=1, \dots, p}$ corresponding to the p principal component $\{V_j\}_{j=1, \dots, p}$ is determined, we transform the vector $\Phi(x) \in \mathbb{R}^l$ to the $\hat{\Phi}(x) \in \mathbb{R}^p$ vector that belongs to the space generated by $\{\Phi(x_b^{(j)})\}_{j=1, \dots, p}$ and the proposed reduced model is

$$\tilde{y}_{\text{reduced}}^{(\text{new})} = \sum_{j=1}^p \hat{a}_j \hat{\Phi}(x^{(\text{new})})_j \tag{21}$$

where

$$\hat{\Phi}(x^{(\text{new})})_j = \langle \Phi(x_b^{(j)}), \Phi(x^{(\text{new})}) \rangle \quad \text{for } j = 1, \dots, p \tag{22}$$

and according to the kernel trick (12), the model (21) is

$$\tilde{y}_{\text{reduced}}^{(\text{new})} = \sum_{j=1}^p \hat{a}_j k_j(x^{(\text{new})}) \tag{23}$$

where

$$k_j(x) = k(x_b^{(j)}, x) \quad \text{for } j = 1, \dots, p \tag{24}$$

The model (23) is less complicate than that provided by the KPCA. The identification problem can be formulated as a minimization of the regularized least square written as:

$$J_r(\hat{a}) = \frac{1}{2} \sum_{i=1}^M \left(y^{(i)} - \sum_{j=1}^p \hat{a}_j k_j(x^{(i)}) \right)^2 + \frac{\rho}{2} \|\hat{a}\|^2 \tag{25}$$

where ρ is a regularization parameter and $\hat{a} = (\hat{a}_1, \dots, \hat{a}_p)^T$ is the parameter estimate vector.

The solution of the problem (25) is

$$\hat{a}^* = (F + \rho I_p)^{-1} G \tag{26}$$

With:

$$G = \begin{pmatrix} \sum_{i=1}^M k_1(x^{(i)}) y^{(i)} \\ \vdots \\ \sum_{i=1}^M k_b(x^{(i)}) y^{(i)} \end{pmatrix} \in \mathbb{R}^p$$

and

$$F = \begin{pmatrix} \sum_{i=1}^M k_1(x^{(i)}) k_1(x^{(i)}) & \dots & \sum_{i=1}^M k_1(x^{(i)}) k_b(x^{(i)}) \\ \vdots & & \vdots \\ \sum_{i=1}^M k_b(x^{(i)}) k_1(x^{(i)}) & \dots & \sum_{i=1}^M k_b(x^{(i)}) k_b(x^{(i)}) \end{pmatrix} \in \mathbb{R}^{p \times p} \tag{27}$$

And $I_p \in \mathbb{R}^{p \times p}$ is the p identity matrix.

The RKPCA algorithm is summarized by the five following steps:

1. Determine the nonzero eigenvalues $\{\mu_j\}_{j=1,\dots,l}$ and the eigenvectors $\{\beta_j\}_{j=1,\dots,l}$ of Gram matrix K .
2. Order the $\{\beta_j\}_{j=1,\dots,l}$ on the decreasing way with respect to the corresponding eigenvalues.
3. For the p retained principal components, choose the $\{x_b^{(j)}\}_{j=1,\dots,p}$ that satisfy (20).
4. Solving (25) to determine $\hat{a}^* \in \mathbb{R}^p$.
5. The reduced RKHS model is given by (23).

5 Online RKPCA method

In this section, we propose an online Reduced Kernel Principal Component Analysis method, which consists on updating the vectors that approximate the principal components. This proposed method is detailed in the following steps.

First step, to determine the optimal value of the parameters μ of the kernel associated to the RKHS model, an offline learning step on an n -observation set $I = \{(x^{(1)}, y^{(1)}), \dots, (x^{(n)}, y^{(n)})\}$ is carried out till the provided RKHS model approximates correctly the nonlinear system. Then, we apply the RKPCA method to reduce the number of the model parameters and the resulting model is written as:

$$\tilde{y}(x) = \sum_{j=1}^p \hat{a}_j k(x_b^{(j)}, x) \tag{28}$$

Let I_n be the set of observations that correspond to the retained principal components. $I_n = \{x_b^{(j)}\}_{j=1,\dots,p}$

At time instant $(n + 1)$, the RKHS model output is obtained according to (28) as:

$$\tilde{y}^{(n+1)} = \sum_{j=1}^p \hat{a}_j k(x_b^{(j)}, x^{(n+1)}) \tag{29}$$

The error between the estimated output and the real one is

$$e^{(n+1)} = |\tilde{y}^{(n+1)} - y^{(n+1)}| \tag{30}$$

If $e^{(n+1)} < \varepsilon_1$, where ε_1 is a given threshold, the model approaches sufficiently the system behavior. Else an update of the RKHS model is required which can be accomplished either by actualizing the model parameters or by actualizing the retained principal components.

In both cases, we calculate the projection of $\Phi(x^{(n+1)})$ on the space F_{kpca} spanned by $\{\Phi(x_b^{(i)})\}_{i=1,\dots,p}$. This projection is denoted $\hat{\Phi}(x^{(n+1)})$ so that its j 'th component is given by:

$$\hat{\Phi}(x^{(n+1)})_j = \langle \Phi(x_b^{(j)}), \Phi(x^{(n+1)}) \rangle = k(x_b^{(j)}, x^{(n+1)}), \tag{31}$$

$j = 1, \dots, p$

A good approximation of $\Phi(x^{(n+1)})$ by $\hat{\Phi}(x^{(n+1)})$ requires satisfying the following condition:

$$\left| \|\hat{\Phi}(x^{(n+1)})\| - \|\Phi(x^{(n+1)})\| \right| < \varepsilon_2 \tag{32}$$

The set I_n is updated to $I_{n+1} = \{I_n, x^{(n+1)}\}$ to determine the parameters $\{\hat{a}_j\}_{j=1,\dots,p}$ of the RKHS model (28).

If (32) isn't satisfied, we actualize the set $\{x_b^{(j)}\}_{j=1,\dots,p}$ using the observation set I_{n+1} then we built the Gram matrix corresponding to I_{n+1} given by:

$$G_{n+1} = \begin{pmatrix} k(x_b^{(1)}, x_b^{(1)}) & \dots & k(x_b^{(1)}, x^{(n+1)}) \\ \vdots & \ddots & \vdots \\ k(x^{(n+1)}, x_b^{(1)}) & \dots & k(x^{(n+1)}, x^{(n+1)}) \end{pmatrix} \quad (33)$$

and we compute its eigenvalues.

According to relations (15), (16) and (33), we determine the new p' principal component. Then, we used the RKPCA to determine the new set $\{x_b^{(j)}\}_{j=1,\dots,p'}$ that approaches the p' retained principal components. The RKPCA model is given by:

$$\tilde{y} = \sum_{j=1}^{p'} \hat{a}_j k(x_b^{(j)}, x) \quad (34)$$

Finally, we estimate the parameters $\hat{a}_j; j = 1, \dots, p'$

In the following, we summarize the algorithm of the online RKPCA method.

6 Online RKPCA algorithm

6.1 Offline phase

1. According to (15) and (16), we determine the p retained principal components resulting from the processing of an n measurement set. Then, we determine the $I_n = \{x_b^{(j)}\}_{j=1,\dots,p}$ set according to (20). The RKHS model is given by:

$$\tilde{y} = \sum_{j=1}^p \hat{a}_j k(x_b^{(j)}, x)$$

6.2 Online phase

1. At time instant $(n + 1)$, we have a new data (x^{n+1}, y^{n+1}) , if $e^{(n+1)} < \varepsilon_1$, the model approaches sufficiently the behavior of the system, else we need to update the RKHS model (28) by the projection $\hat{\Phi}(x^{n+1})$ given by (31).
2. If (32) is satisfied, we use the set $I_{n+1} = \{I_n, x^{(n+1)}\}$ to actualize the parameters $\{\hat{a}_j\}_{j=1,\dots,p}$, else we update the $\{x_b^{(j)}\}_{j=1,\dots,p}$ set using the I_{n+1} set and the relations (15), (16) and (33). The new RKHS model is given by (34), and the a_j parameters of the model can be determined using the least square method.

7 Simulations

The proposed method has been tested for modeling a Wiener Hammerstein benchmark and a chemical reactor.

7.1 Description of wiener Hammerstein benchmark model

The system to be modelled is sketched by Fig. 1. It consists on an electronic nonlinear system with a Wiener Hammerstein structure that was built by Gerd Vendesteen [21]. This process was adopted as a nonlinear system benchmark in SYSID 2009.

7.2 Results

To build the RKHS model, we use the RBF Kernel (Radial Basis Function)

$$K(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\mu^2}\right), \quad \mu = 88 \quad (35)$$

We use a heuristic approach to select the input vector that yields the minimal normalized mean square error between real output and estimated one. This approach is called sequential forward search, in which each input is selected sequentially. The selected vector is: $x(k) = \{u(k - 1), u(k - 2), u(k - 4), \dots, u(k - 15), y(k - 1)\}^T \in \mathbb{R}^{15}$ selected with validation Normalized Mean Square Error NMSE of 0.063.

The chosen thresholds are

$$\varepsilon_1 = 0.09, \quad \varepsilon_2 = 0.01$$

We performed the online identification using the online RKPCA algorithm developed in Sect. 5. The total number of observations is 187,000.

The parameter estimate vector is

$$\hat{a} = 10^4 \times \begin{bmatrix} -1.95 & -0.05 & 1.62 & -0.06 & 0.43 & 0.25 \\ -0.24 & 0.74 & 0.02 & -0.88 & 0.20 & -1.3 & 0.96 \\ -0.09 & 0.52 & -0.17 \end{bmatrix}^T \in \mathbb{R}^{16}$$

The number of the retained principal components is $p' = 16$. They form the columns of the following matrix.

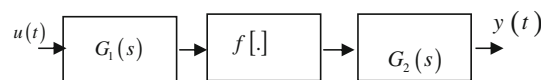


Fig. 1 Wiener Hammerstein benchmark

$$x_b = \begin{pmatrix} 0.1548 & -0.0130 & 0.5902 & 0.4587 & -1.8770 & -0.3059 & -1.2271 & -0.5081 & -0.0058 & 0.5579 & -0.0031 & -0.6057 & -0.1593 & -0.1160 & 0.4391 & -0.0076 \\ 0.3063 & 0.0124 & 0.5394 & 1.0115 & -1.2494 & -0.3609 & -1.5622 & -1.1670 & -0.0052 & 0.6482 & -0.0076 & -1.2271 & -0.1878 & -0.0141 & 0.1023 & -0.0024 \\ 0.4549 & -0.6057 & 0.1463 & 1.3868 & -0.2050 & -0.3986 & -1.6233 & -1.3631 & -0.0076 & 0.4391 & -0.0038 & -1.5622 & -0.0165 & 0 & -0.1593 & -0.0130 \\ 0.6884 & -1.2271 & -0.3183 & 1.2855 & 0.7722 & -0.3921 & -1.4379 & -1.1622 & -0.0024 & 0.1023 & -0.0076 & -1.6233 & 0.2627 & -0.0096 & -0.1878 & 0.0124 \\ 1.0424 & -1.5622 & -0.4996 & 0.7018 & 1.2917 & -0.1648 & -1.1519 & -0.8728 & -0.0130 & -0.1593 & -0.0041 & -1.4379 & 0.5648 & -0.0024 & -0.0165 & -0.6057 \\ 1.4352 & -1.6233 & -0.2486 & -0.0577 & 1.2196 & 0.3866 & -0.7533 & -0.7543 & 0.0124 & -0.1878 & -0.0072 & -1.1519 & 0.8893 & -0.0082 & 0.2627 & -1.2271 \\ 1.6841 & -1.4379 & 0.2949 & -0.5806 & 0.6960 & 1.0159 & -0.2881 & -0.8175 & -0.6057 & -0.0165 & -0.0048 & -0.7533 & 1.1832 & -0.0031 & 0.5648 & -1.5622 \\ 1.6209 & -1.1519 & 0.8508 & -0.6249 & 0.0398 & 1.1821 & 0.2074 & -0.8614 & -1.2271 & 0.2627 & -0.0065 & -0.2881 & 1.2989 & -0.0076 & 0.8893 & -1.6233 \\ 1.1935 & -0.7533 & 1.2003 & -0.2688 & -0.4113 & 0.5181 & 0.5579 & -0.6887 & -1.5622 & 0.5648 & -0.0048 & 0.2074 & 1.0348 & -0.0038 & 1.1832 & -1.4379 \\ 0.5020 & -0.2881 & 1.2693 & 0.1535 & -0.4522 & -0.7251 & 0.6482 & -0.2781 & -1.6233 & 0.8893 & -0.0062 & 0.5579 & 0.3567 & -0.0076 & 1.2989 & -1.1519 \\ -0.2245 & 0.2074 & 1.0915 & 0.2609 & -0.0752 & -1.7504 & 0.4391 & 0.1919 & -1.4379 & 1.1832 & -0.0048 & 0.6482 & -0.5081 & -0.0041 & 1.0348 & -0.7533 \\ -0.7056 & 0.5579 & 0.7615 & -0.1370 & 0.5456 & -1.8503 & 0.1023 & 0.4755 & -1.1519 & 1.2989 & -0.0065 & 0.4391 & -1.1670 & -0.0072 & 0.3567 & -0.2881 \\ 0.4249 & 0.6157 & -0.1363 & 1.4868 & 0.2250 & -0.3786 & 1.5243 & -1.2731 & -0.0076 & 0.4431 & -0.0048 & -1.5421 & -0.0187 & -0.0045 & -0.1373 & -0.0172 \\ -0.3052 & 0.4391 & 0.2393 & -1.6206 & 1.4795 & 0.0906 & -0.1878 & 0.1827 & -0.2881 & 0.3567 & -0.0058 & -0.1593 & -1.1622 & -0.0065 & -1.1670 & 0.5579 \\ 0.1257 & -0.3746 & 0.0989 & 0.3945 & -0.2572 & -0.2520 & -0.5236 & -0.1092 & -0.1274 & -0.2397 & -0.0666 & -0.4865 & 0.0670 & -0.1569 & -0.0841 & -0.2417 \end{pmatrix}$$

In Fig. 2a–c, we present the system and the model output during the online identification. We picked the performance of our algorithm in the training sample windows [4800 5800], $[10^5 \ 1.110^5]$ and [183200 184200]. We remark that the model output is in concordance with the system output, indeed the Normalized mean Square Error is equal to 0.087. This shows the good performances of the proposed online identification method.

To evaluate the performance of the proposed method, we plot in Fig. 3, the evolution of the NMSE. We notice that the error goes down when the number of observation goes high.

In the Table 1, we summarized the results of the online identification algorithm in terms of kernel parameter, NMSE, and parameters number.

7.3 Chemical reactor modeling

7.3.1 Process description

The process is a Continuous Stirred Tank Reactor CSTR which is a nonlinear system used for the conduct of the chemical reactions [8]. A diagram of the reactor is given in Fig. 4.

The physical equations describing the process are

$$\begin{aligned} \frac{dh(t)}{dt} &= w_1(t) + w_2(t) - 0, 2\sqrt{h(t)} \\ \frac{dC_b(t)}{dt} &= C_{b1} - C_b(t)\frac{w_1(t)}{h(t)} + (C_{b2} - C_b(t))\frac{w_2(t)}{h(t)} \\ &\quad - \frac{k_1 \cdot C_b(t)}{(1 + k_2 \cdot C_b(t))^2} \end{aligned} \quad (36)$$

where $h(t)$ is the height of the mixture in the reactor of the feed of reactant 1 w_1 (resp, reactant 2, w_2) with concentration C_{b1} (resp. C_{b2}). The feed of product of the reaction is w_0 and its concentration is C_b . k_1 and k_2 are consuming reactant rate. The temperature in the reactor is assumed constant and equal to the ambient temperature. We are interested by modeling the subsystem presented in Fig. 5.

For the purpose of the simulations, we used the CSTR model of the reactor provided with Simulink of Matlab. The parameter $\mu = 208$ is determine using the cross-validation technique in the offline phase.

The input vector of RKHS model is

$$x(k) = [w_1(k-1), w_1(k-2), cb(k-1), cb(k-2)]^T \quad (37)$$

The number of observations is 300.

In Fig. 6, we represent the online reduced kernel principal component analysis output as well as the system output. We notice concordance between both outputs with a Normalized mean Square Error equal to 0.083%.

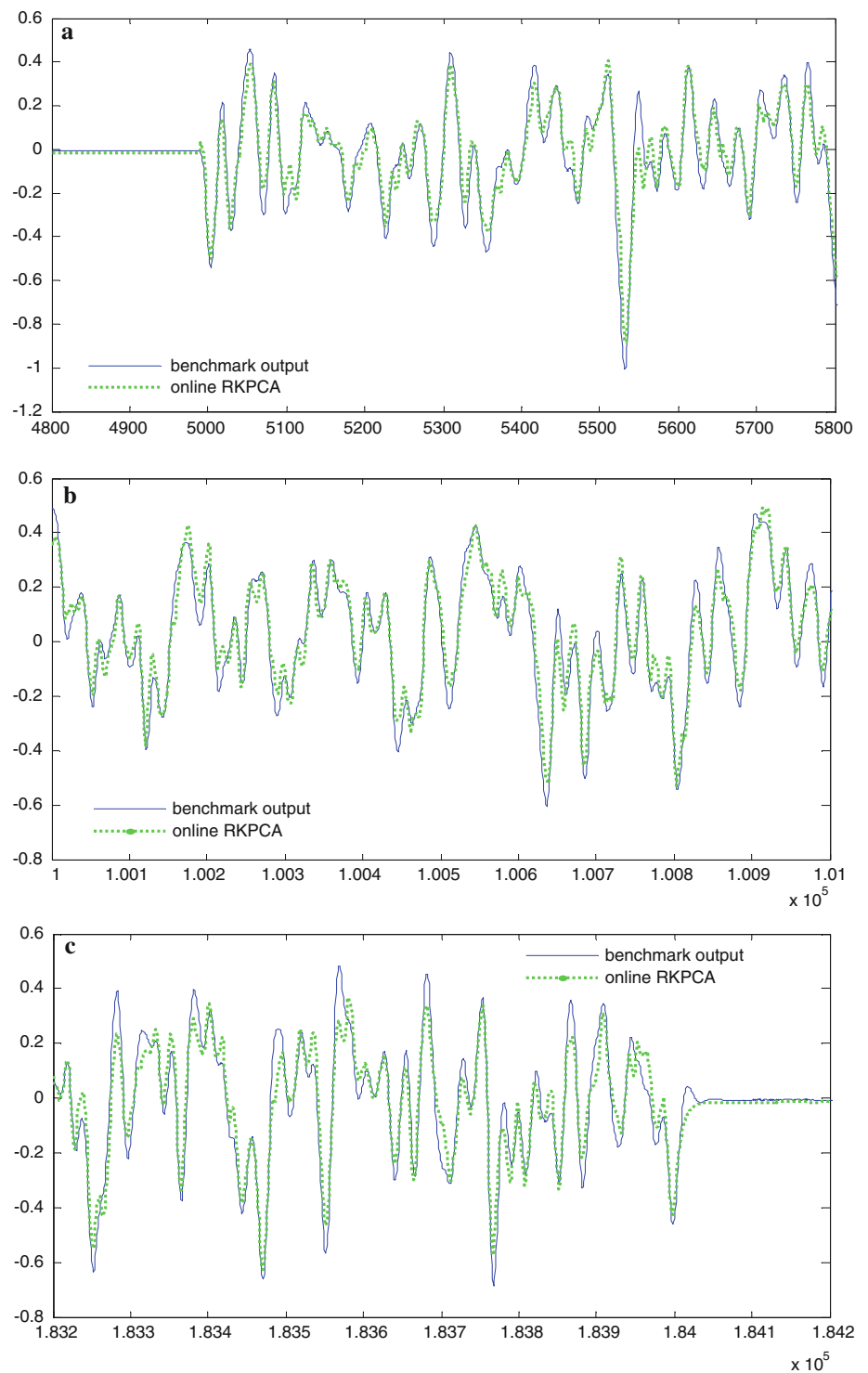
In Fig. 7, we draw the evolution of the NMSE, we notice that the online identification algorithm presents readily an error less than 1% since the 20th training sample.

In the Table 2, we summarized the performance of the online identification algorithm in terms of kernel parameter, NMSE, and parameters number.

8 Conclusion

In this paper, we have proposed an online reduced kernel principal component analysis method for nonlinear system parameter identification. Through several experiments, we

Fig. 2 System and model outputs during the online identification



showed the accuracy and good scaling properties of the proposed method. This algorithm has been tested for identifying a Wiener Hammerstein benchmark model and a

chemical reactor, and the results were satisfying. The proposed technique may be very helpful to design an adaptive control strategy of nonlinear systems.

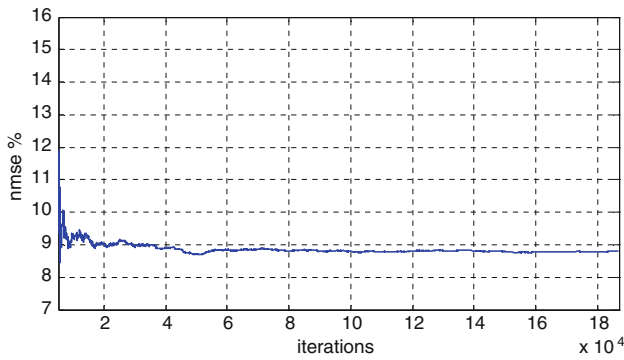


Fig. 3 Evolution of NMSE

Table 1 Results of the Wiener Hammerstein benchmark identification with the online RKPCA algorithm

| Kernel parameter | NMSE | Observation number | Number of the model parameters |
|------------------|-------|--------------------|--------------------------------|
| RBF, $\mu = 88$ | 0.087 | 187,000 | 16 |

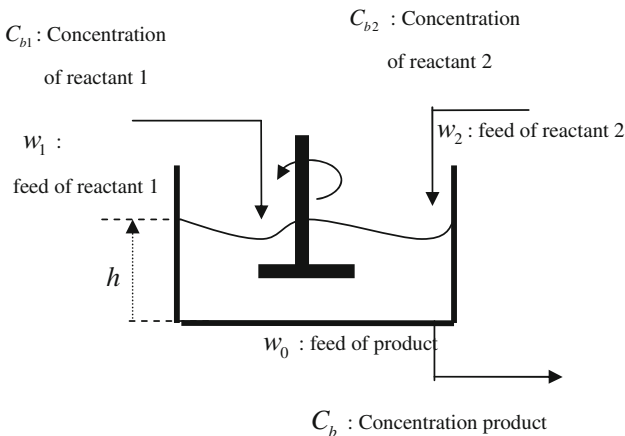


Fig. 4 Chemical reactor diagram

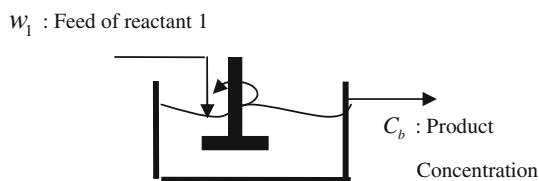


Fig. 5 Considered subsystem

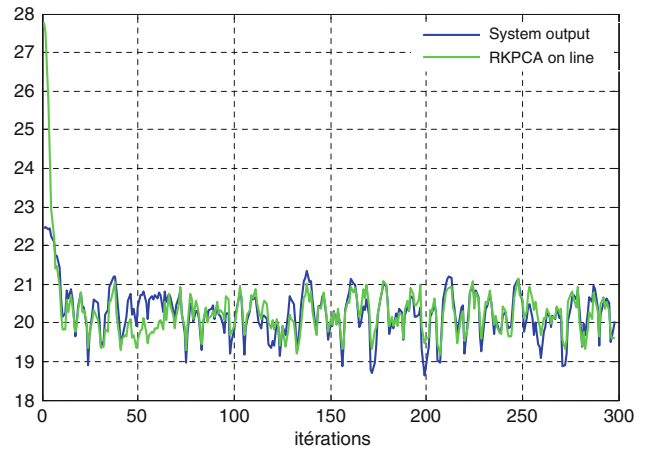


Fig. 6 System and model outputs during the online identification

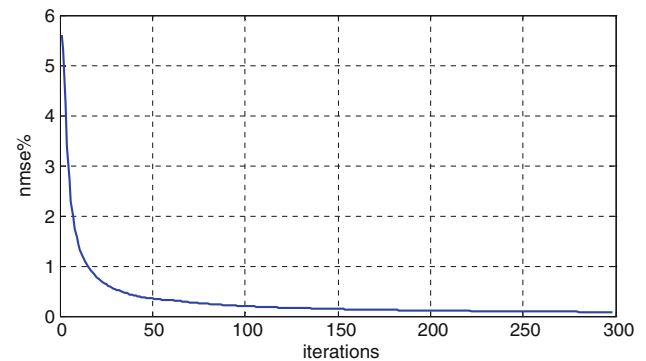


Fig. 7 Evolution of NMSE

Table 2 Results of the chemical reactor identification with the online RKPCA algorithm

| Kernel parameter | NMSE % | Observation number | Number of the model parameters |
|------------------|--------|--------------------|--------------------------------|
| RBF, $\mu = 208$ | 0.083 | 300 | 15 |

References

1. Aissi I (2009) Modélisation, Identification et Commande Prédictive des systèmes non linéaires par utilisations des espaces RKHS, thèse de doctorat, ENIT Tunisie
2. Scholkopf B, Smola A, Muller K-R (1998) Nonlinear component analysis as kernel eigenvalue problem. *Neural Comput* 10:1299–1319
3. Scholkopf B, Smola A (2002) *Learning with kernels*. The MIT press, Cambridge
4. Richard C, Bermudez JCM, Honeine P (2009) Online prediction of time series data with kernels. *IEEE Trans Signal Process* 57:1058–1067

5. Kuzmin D, Warmuth MK (2007) Online Kernel PCA with entropic matrix updates. In: Proceedings of the 24 the international conference on machine learning, Corvallis, OR
6. Mercer J (1909) Functions of positive and negative type and their connection with the theory of integral equations. *Philos Trans R Soc* 209:415–446
7. Suykens JAK, Van Gestel T, De Brabanter J, De Moor B, Vandewalle J (2002) Least squares support vector machines. World Scientific, Singapore
8. Demuth H, Beale M, Hagan M (2007) Neural network toolbox 5, User's guide, The MathWorks
9. Wahba G (2000) An introduction to model building with reproducing Kernel Hilbert spaces. Technical report No 1020, Department of Statistics, University of Wisconsin-Madison
10. Aronszajn N (1950) Theory of reproducing kernels. *Trans Am Math Soc* 68(3):337–404
11. Cristianini N, Shawe-Taylor J (2000) An introduction to support vector machines. Cambridge University Press, Cambridge
12. Rosipal R, Trecho LJ (2001) Kernel partial least squares in reproducing Kernel Hilbert spaces. *J Mach Learn Res* 2:97–123
13. Fernandez R (1999) Predicting time series with a local support vector regression machine. In: ACAI conference
14. Gunter Simon, Schraudolph NN, Vishwanathan SVN (2007) Fast iterative kernel principal component analysis. *J Mach Learn Res* 8:1893–1918
15. Chin T-J, Schindler K, Suter D (2006) Incremental Kernel SVD for face recognition with image sets. In: Proceedings of the 7th international conference on automatic face and gesture recognition FGR
16. Lauer F (2008) Machines à Vecteurs de Support et Identification de Systèmes Hybrides, thèse de doctorat de l'Université Henri Poincaré—Nancy 1, octobre 2008, Nancy, France
17. Taouali O, Aissi I, Villa N, Messaoud H (2009) Identification of nonlinear multivariable processes modelled on reproducing Kernel Hilbert space: application to tennessee process. In: ICONS, 2nd IFAC international conference on intelligent control systems and signal processing. Istanbul, pp 1–6
18. Vovk V (2008) Leading strategies in competitive on-line prediction. *Theor Comput Sci* 405:285–296
19. Vapnik VN (1995) The nature of statistical learning theory. Springer, New York
20. Vapnik VN (1998) Statistical learning theory. Wiley, New York
21. Vandersteen G (1997) Identification of linear and nonlinear systems in an errors-in-variables least squares and total least squares framework, PhD thesis, Vrije Universiteit Brussel
22. Vovk V (2006) On-line regression competitive with reproducing kernel Hilbert spaces, Technical report arXiv:cs.LG/0511058 (version 2), arXiv.org e-Print archive, Jan 2006
23. Veropoulos K, Cristianini N, Campbell C (1999) The application of support vector machines to medical decision support: a case study, ACAI conference
24. Ghatte VN, Dudul SV (2009) Induction machine fault detection using support vector machine based classifier. *WSEAS Trans Syst* 8:591–603
25. Wanga W, Mena C, Lub W (2008) Online prediction model based on support vector machine. *Neurocomputing* 71:550–558
26. Yang MH (2002) Kernel eigenfaces vs. Kernel first faces: face recognition using kernel method, In: IEEE FGR, pp 215–220