Adaptive identification of time delays in nonlinear dynamical models

Huanfei Ma,[1,2] Bing Xu,[1] Wei Lin,[1,3,*] and Jianfeng Feng[1,4]
[1]*School of Mathematical Sciences and Centre for Computational Systems Biology, Fudan University, Shanghai 200433, China*
[2]*School of Mathematical Sciences, Soochow University, Suzhou 215006, China*
[3]*CAS-MPG Partner Institute for Computational Biology, Shanghai 200031, China*
[4]*Department of Computer Science, University of Warwick, Coventry CV47AL, United Kingdom*
(Received 10 November 2009; revised manuscript received 21 October 2010; published 20 December 2010)

This paper develops an adaptive synchronization strategy to identify both discrete and distributed time delays in nonlinear dynamical models. In contrast with adaptive techniques for parameter estimation in the literature, the adaptive strategy developed here for time-delay identification invites more precise results that have physical and dynamical importance. It is analytically and numerically found that distributed time delays in a model with an asymptotically stable steady state can be adaptively identified, and which is different from the case of discrete time-delays identification. Other aspects of the strategy developed here, for time-delay identification, are illustrated by several representative dynamical models. Aside from illustrations for toy models and their generated data, the strategy developed is used with experimental data, to identify a time delay, called transcriptional delay, in a model describing the transcription of messenger RNAs (mRNAs) for Notch signaling molecules.

DOI: 10.1103/PhysRevE.82.066210 PACS number(s): 05.45.−a, 02.30.Ks, 87.19.ln

I. INTRODUCTION

Two research directions for nonlinear dynamical systems or systems biology have been previously suggested [1]: one involves data-driven research, and the other focuses on model-driven investigations. These two directions are not isolated, however, but are connected. For example, in integrated investigations, qualitative consistency is expected between the real data and the data produced by an *a priori* dynamical model with appropriate configurations [2,3]. Among all the dynamical models, the model, with a set of time delays, may be much closer to the real system and capable of describing richer behavior including physical, chemical and circadian oscillations [4–6]. Given a collection of real data and an *a priori* dynamical model with unknown time delays, several questions then arise. The first question is: "How to develop a strategy for an accurate identification of time delays with the data produced by the *a priori* dynamical model that is supposed to be capable of capturing the real system exactly?" The subsequent questions, once the developed strategy is theoretically justified, naturally become: "Is the specific strategy suitable also for coping with real data?" and "Is it robust against a certain amount of perturbation to the systems and the parameters?" One may also be: "Is there any intrinsic difference between parameter estimation and time-delay identification when a specific strategy is used?"

Attempts to answer the above questions have been made at both theoretical and experimental levels, and various algorithms and strategies for time-delay identification have been developed in the literature [7–12]. On the face of it, identification of time delays could be seen as estimation of a particular kind of parameter; nevertheless, it is not easy because (i) the models associated with identification algorithms generally consist of functional differential equations, and have the property of being infinite dimensional [13], and (ii) time delays have diverse forms, such as the discrete time delays, distributed time delays, finite or infinite length time delays, and time-varying delays [13–15].

As a matter of fact, the majority of existing works treats time-delay identification as parameter estimation, and consider the discrete time delays only, and rarely take into account the infinite-dimension property of the models. For instance, some of the existing algorithms are suitable only for the linear or quasilinear models with discrete time delays [7,8]. Some of the algorithms require external perturbations to internal state variables and then need to take account of finite peaks that emerge in the plots of correlation functions [9,10]. These algorithms are not appropriate for identifying either the distributed or time-varying delays since the emerging infinite peaks, or location-varying peaks, may lead the counting method to fail. Some of the algorithms cannot be implemented adaptively and are, in general, computationally expensive with a complexity $O(L^d)$ [11], where L stands for the length of the observed time series and d the number of discrete steps in the interval that contains a single unknown time delay. Some of the algorithms still need a proper choice of coupling gain and consider neither distributed time delays nor some intrinsic characters and restrictions of time-delay identification [12].

In this paper, we identify time delays in nonlinear dynamical models by using an adaptive synchronization strategy. This strategy, making a good use of chaos synchronization, was initially proposed to estimate parameters in chaotic dynamical systems [16,17]. Recently, its advantages and limitations for parameter estimation in any nonlinear dynamical systems have been systematically investigated by virtue of dynamical systems theory [18,19]. However, due to the property of infinite dimensionality and the diverse forms of time delays, time-delay identification, through an adaptive strategy, is intrinsically different from parameter estimations in systems described by ordinary differential equations or

*FAX: 86-21-65646073; wlin@fudan.edu.cn

even by functional differential equations. Indeed, investigations presented in this paper generate more significant results. These indicate several advantages over the existing theories including: (i) identification of distributed time delays, (ii) online identification of time-varying delays that have lower fluctuating frequencies, (iii) simultaneous identification of time delays and parameters within a unified framework, (iv) differentiation of the specific forms of the time delays that can be identified adaptively, and (v) robustness against a certain amount of perturbations in coping with real data.

The remainder of this paper is organized as follows. In Sec. II, we introduce a model that is supposed to be capable of describing a real system exactly, and then propose an adaptive synchronization strategy to identify the time delays in this model. Also in this section, we validate the proposed strategy by the theory of functional differential equations. We include in Appendices containing more technical results and arguments on the identifiability of a given time-delayed model. In Sec. III, we intensively test the proposed strategy by several typical time-delay models, illustrating the theoretical results established in the paper. Still, in this section, we numerically verify the robustness of the proposed strategy against perturbations to parameters and system. In Sec. IV, with experimental data, we identify a time delay, called transcriptional delay, in a model describing the transcription of mRNAs for Notch signaling molecules. Finally, we end the paper with discussions and concluding remarks.

II. ADAPTIVE SYNCHRONIZATION STRATEGY FOR IDENTIFYING TIME DELAYS

Model and Strategy Description. To begin, consider a general time-delayed model described by the following functional differential equations:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{F}(\boldsymbol{x}_t), \quad t \geq 0, \tag{1}$$

with the initial condition $\boldsymbol{x}(\theta)=\boldsymbol{\psi}(\theta)$ in which $\boldsymbol{\psi} \in \mathcal{C} \triangleq \mathbf{C}([-\tau_M, 0], \mathbb{R}^n)$. Here, $\tau_M > 0$ is a constant sufficiently large, $\mathbf{C}([-\tau_M, 0], \mathbb{R}^n)$ represents the family of all continuous $\mathbb{R}^n$-valued functions on the interval $[-\tau_M, 0]$, the state vector $\boldsymbol{x}(t)=[x_1(t), x_2(t), \ldots, x_n(t)]^{\mathrm{T}} \in \mathbb{R}^n$, and the distributed form of the state vector $\boldsymbol{x}_t(\theta)=\boldsymbol{x}(t+\theta)$ for $\theta \in [-\tau_M, 0]$. Moreover, each component of the vector field $\boldsymbol{F}(\boldsymbol{\varphi}) = [F_1(\boldsymbol{\varphi}), \ldots, F_n(\boldsymbol{\varphi})]^{\mathrm{T}}$ is assumed to be in the following explicit form:

$$F_i(\boldsymbol{\varphi}) = \sum_{j=1}^{k_i} f_{ij}(\boldsymbol{\varphi}(-\tau_{ij})) + \sum_{j=k_i+1}^{p_i} \int_{-\tau_{ij}}^{0} f_{ij}(\boldsymbol{\varphi}(\theta)) d\theta, \tag{2}$$

where k_i and p_i are finite natural numbers with $k_i < p_i$, $\boldsymbol{\varphi} \in \mathcal{C}$, each function f_{ij} at least satisfies the local Lipschitz condition, and each time-delay constant $\tau_{ij} \in [0, \tau_M]$. Furthermore, to avoid the trivial case of no time delays, it is assumed that there exists at least one pair of indexes (i_0, j_0) satisfying $\tau_{i_0 j_0} \neq 0$. Let $\boldsymbol{\varphi}(\theta)=\boldsymbol{x}_t(\theta)$. Then, each component of the vector field becomes

$$F_i(\boldsymbol{x}_t) = \sum_{j=1}^{k_i} f_{ij}(\boldsymbol{x}(t-\tau_{ij})) + \sum_{j=k_i+1}^{p_i} \int_{t-\tau_{ij}}^{t} f_{ij}(\boldsymbol{x}(s)) ds,$$

where the first summation term corresponds to the case of discrete time delays and the second term to the case of distributed time delays. Mathematically, the vector field [Eq. (2)] can be rewritten in a more compact form (see Appendix A). Though the above settings cannot cover all time-delayed systems, numerous well-known dynamical models with time delays can be expressed in the form of Eq. (2). Representative examples include the delayed Logistic model, the Mackey-Glass model, the artificial neural network model with discrete and distributed time delays, the predator-prey model with distributed time delays, the gene transcription model with discrete time delays, and so on.

Assume that the state vector $\boldsymbol{x}(t)=\boldsymbol{x}(t;\boldsymbol{\psi})$ represents the data generated by a real system, and that model (1) with unknown time delays can capture the system exactly. Now, with the data $\boldsymbol{x}(t)$, the objective is to design a strategy, including a coupling model, dynamical estimators, and time-variant control gains, for identifying those unknown time delays automatically. It is expected that once model (1) and the coupling model approach complete synchronization, the estimators converge to the true values of the unknown time delays. For this purpose, the coupling model is designed as

$$\dot{\boldsymbol{y}}(t) = \hat{\boldsymbol{F}}(\boldsymbol{y}_t) + \boldsymbol{\epsilon}(t) \circledast [\boldsymbol{y}(t) - \boldsymbol{x}(t)], \tag{3}$$

where the coupling term $\boldsymbol{\epsilon}(t) \circledast [\boldsymbol{y}(t)-\boldsymbol{x}(t)] = \{\epsilon_1(t)[y_1(t) - x_1(t)], \cdots, \epsilon_n(t)[y_n(t)-x_n(t)]\}^{\mathrm{T}}$, and the vector field $\hat{\boldsymbol{F}}(\boldsymbol{\varphi}) = [\hat{F}_1(\boldsymbol{\varphi}), \ldots, \hat{F}_n(\boldsymbol{\varphi})]^{\mathrm{T}}$ has the same form of $\boldsymbol{F}$ in Eq. (2) except that all the constants τ_{ij} are replaced with the variables $\nu_{ij}(t)$, that is

$$\hat{F}_i(\boldsymbol{\varphi}) = \sum_{j=1}^{k_i} f_{ij}(\boldsymbol{\varphi}(-\nu_{ij})) + \sum_{j=k_i+1}^{p_i} \int_{-\nu_{ij}}^{0} f_{ij}(\boldsymbol{\varphi}(\theta)) d\theta,$$

Here, the estimators of the time delays, $\boldsymbol{\nu}_t=\{\nu_{ij_t}\}$, and the control gains of the couplings, $\boldsymbol{\epsilon}_t=[\epsilon_{1_t}, \ldots, \epsilon_{n_t}]^{\mathrm{T}}$, are designed to obey the following adaptive rules, respectively,

$$\dot{\nu}_{ij}(t) = -\delta_{ij}[y_i(t) - x_i(t)] \frac{\partial \hat{F}_i}{\partial \nu_{ij}},$$

$$\dot{\epsilon}_i(t) = -r_i[y_i(t) - x_i(t)]^2,$$

$$i = 1, 2, \ldots, n, \quad j = 1, 2, \ldots, p_i, \tag{4}$$

where the positive constants δ_{ij} and r_i are adjustable in accordance with the need of computational accuracy. The initial conditions for coupling model (3) with the adaptive rules [Eq. (4)] are selected as $\boldsymbol{y}(\theta)=\boldsymbol{\eta}(\theta)$, $\boldsymbol{\nu}(\theta)=\boldsymbol{\xi}(\theta)$, and $\boldsymbol{\epsilon}(\theta)=\boldsymbol{\zeta}(\theta)$, in which $\boldsymbol{\eta}, \boldsymbol{\zeta} \in \mathcal{C}$ and each row of $\boldsymbol{\xi}$ satisfies $\boldsymbol{\xi}_i \in \mathcal{C}_{p_i} \triangleq \mathbf{C}([-\tau_M, 0], \mathbb{R}^{p_i})$ for all i.

Convergence validation. In what follows, we theoretically validate the effectiveness of the proposed strategy for all bounded trajectories that are generated by coupled models (1) and (3) with the adaptive rules [Eq. (4)]. For a specific bounded trajectory $[\boldsymbol{x}_t, \boldsymbol{y}_t, \boldsymbol{\epsilon}_t, \boldsymbol{\nu}_t]$ with $\nu_{ij_t} \in [0, \tau_M]$, we have

the existence of $\lim_{t\to\infty}\boldsymbol{\epsilon}(t)$ because of the monotonicity and boundedness of each component of the control gain variable $\boldsymbol{\epsilon}(t)$. Thus an integration of this control gain gives

$$\epsilon_i(0)-\epsilon_i(+\infty)=r_i\int_0^{+\infty}[y_i(s)-x_i(s)]^2ds<+\infty,$$

which implies that the square of the error $\boldsymbol{e}(t)=\boldsymbol{y}(t)-\boldsymbol{x}(t)$ is integrable. Furthermore, the error satisfies

$$\dot{\boldsymbol{e}}(t)=\dot{\boldsymbol{y}}(t)-\dot{\boldsymbol{x}}(t)=\hat{\boldsymbol{F}}(\boldsymbol{y}_t)+\boldsymbol{\epsilon}(t)\circledast[\boldsymbol{y}(t)-\boldsymbol{x}(t)]-\boldsymbol{F}(\boldsymbol{x}_t). \tag{5}$$

Hence, $\dot{\boldsymbol{e}}(t)$ is uniformly bounded due to the local Lipschitz property of the vector field and the boundedness of the trajectory. This uniform boundedness, along with the integrability and continuity of $\boldsymbol{e}(t)$, implies $\boldsymbol{e}(t)\to\boldsymbol{0}$ as $t\to+\infty$.

Moreover, since coupled models (1) and (3) with the adaptive rules [Eq. (4)] constitute an autonomous system and have bounded trajectories as assumed above, the ω-limit set, denoted by

$$\Omega=\{[\bar{\boldsymbol{\psi}},\bar{\boldsymbol{\eta}},\bar{\boldsymbol{\zeta}},\bar{\boldsymbol{\xi}}]\in\mathcal{C}\times\mathcal{C}\times\mathcal{C}\times\mathcal{C}_p|[\boldsymbol{x}_{t_s}^{\boldsymbol{\psi}},\boldsymbol{y}_{t_s}^{\boldsymbol{\eta}},\boldsymbol{\epsilon}_{t_s}^{\boldsymbol{\zeta}},\boldsymbol{\nu}_{t_s}^{\boldsymbol{\xi}}]$$
$$\to[\bar{\boldsymbol{\psi}},\bar{\boldsymbol{\eta}},\bar{\boldsymbol{\zeta}},\bar{\boldsymbol{\xi}}],\quad \text{as } t_s\to+\infty \text{ and } s\to\infty\},$$

is nonempty, invariant, and connected according to the theory of functional differential equations (see [13] and references therein). Here, $[\boldsymbol{x}_{t_s}^{\boldsymbol{\psi}},\boldsymbol{y}_{t_s}^{\boldsymbol{\eta}},\boldsymbol{\epsilon}_{t_s}^{\boldsymbol{\zeta}},\boldsymbol{\nu}_{t_s}^{\boldsymbol{\xi}}]$ represents the point on the orbit of the coupled model starting from the initial value $[\boldsymbol{\psi},\boldsymbol{\eta},\boldsymbol{\zeta},\boldsymbol{\xi}]$, and the functions space $\mathcal{C}_p$ is the Cartesian product of all $\mathcal{C}_{p_i}$. Notice that $\boldsymbol{e}_t(\theta)\equiv\boldsymbol{0}$ in the ω-limit set. Then, restricted to this limit set, the functions $\bar{\boldsymbol{\zeta}}$ and $\bar{\boldsymbol{\xi}}$ become constants, that is, $\bar{\boldsymbol{\psi}}(\theta)\equiv\bar{\boldsymbol{\eta}}(\theta)$, $\bar{\boldsymbol{\zeta}}(\theta)\equiv\boldsymbol{\zeta}^*$, and $\bar{\boldsymbol{\xi}}(\theta)\equiv\boldsymbol{\xi}^*$ for all θ and some constants $\boldsymbol{\zeta}^*$, $\boldsymbol{\xi}^*$. Hence, from Eq. (5), it follows that in the ω-limit set,

$$\boldsymbol{F}(\boldsymbol{x}_t(\theta))-\hat{\boldsymbol{F}}(\boldsymbol{x}_t(\theta))=\boldsymbol{0}, \tag{6}$$

where $\boldsymbol{x}_t(\theta)$ is the observable synchronized orbit contained in the synchronization manifold for all t. In order to determine whether or not Eq. (6) implies a successful identification of the time delays, we import the following definition on the *identifiability* of a given vector field with time delays.

Definition. For some index i_0, the component F_{i_0} in the vector field $\boldsymbol{F}$ is said to be *identifiable* with respect to $\tau_{i_0 j}$ $(j=1,\ldots,p_{i_0})$ in the set $\mathfrak{M}$, if the equation $F_{i_0}(\boldsymbol{x}_t,\tau_{i_0 1},\ldots,\tau_{i_0 p_{i_0}})=F_{i_0}(\boldsymbol{x}_t,\nu_{i_0 1},\ldots,\nu_{i_0 p_{i_0}})$ for all t and $\boldsymbol{x}_t\in\mathfrak{M}$ implies $\tau_{i_0 j}=\nu_{i_0 j}$ for all $j=1,\ldots,p_{i_0}$. Here, the component is rewritten as $F_i(\boldsymbol{x}_t)=F_i(\boldsymbol{x}_t,\tau_{i1},\ldots,\tau_{ip_i})$ for $i=1,\cdots,n$, and $\mathfrak{M}$ is some subset of $\mathcal{C}$. The vector field $\boldsymbol{F}$ is said to be *identifiable* with respect to τ_{ij} for all $i,j=1,\ldots,p_i$ in the set $\mathfrak{M}$, if each component F_i is identifiable with respect to $\tau_{ij}(j=1,\ldots,p_i)$ in the set $\mathfrak{M}$.

Now, together with the definition and the above-performed arguments, we approach a conclusion: "Assume that coupled models (1) and (3) with the adaptive rules [Eq. (4)] have bounded trajectories. Also assume that the vector field is identifiable on the synchronized manifold. Then the estimators, $\nu_{ij}(t)$, can converge to the real values of those unknown time delays, τ_{ij}, respectively."

We stress that not all vector fields are identifiable with an adaptive strategy. This point of view has been seldom reported in the literature, to the best of our knowledge. As a matter of fact, the determination of identifiability is a topic of theoretical and practical significance, which requires the use of the infinite-dimension property in the ω-limit set. For conciseness, the detailed investigations on the identifiability of representative vector fields are included in Appendix B. Some of the results obtained there are summarized as follows. When a specific model with discrete time delays possesses stable steady states (also called equilibria), its vector field is not definitely identifiable; when this model possesses stable periodic orbits, its vector field is identifiable in some generalized sense. More interestingly, when a model with distributed time delays has either stable steady states or stable periodic orbits or even chaotic attractors, the vector field is identifiable under some mild conditions. Specific examples in the next section further illustrate these results that are explicitly obtained in Appendix B.

We remark that the argument performed above crucially depends on the boundedness assumption on the trajectories of coupled models (1) and (3) with the adaptive rules [Eq. (4)]. Although a verification of this assumption is no easy job for models with general forms, it can be demonstrated for some typical models either theoretically (see Appendix C) or numerically. The examples in the following sections also justify this prior assumption numerically.

We remark also that the assumption that model (1) is capable of capturing the real system exactly is a prerequisite in the above theoretical discussions. In practice, various forms of perturbations are unavoidable in real systems, so that a model with mismatched parameters or established through noise-perturbed data sometimes cannot describe the real system exactly. The examples in the following sections illustrate the robustness of the adaptive strategy proposed here. That is, under a certain amount of perturbation, the data reproduced by the model with the identified time delays is not completely consistent with the real data but is qualitatively consistent.

Furthermore, due to the online property, the proposed adaptive strategy is applicable to identifying the time-varying delays with lower fluctuating frequencies. This advantage is also illustrated by specific examples in the next section.

III. ILLUSTRATIVE EXAMPLES

In this section, several representative time-delayed models are used as toy models to illustrate the feasibility of the above-proposed adaptive synchronization strategy for time-delays identification.

Example 1. Chaotic models with a time-invariant discrete delay and with a time-varying delay. Consider the Mackey-Glass model [20] in the form of

$$\dot{x}=\frac{ax(t-\tau)}{1+[x(t-\tau)]^b}-cx. \tag{7}$$

The Mackey-Glass model was initially introduced as a model for regeneration of blood cells and later became a prototypi-

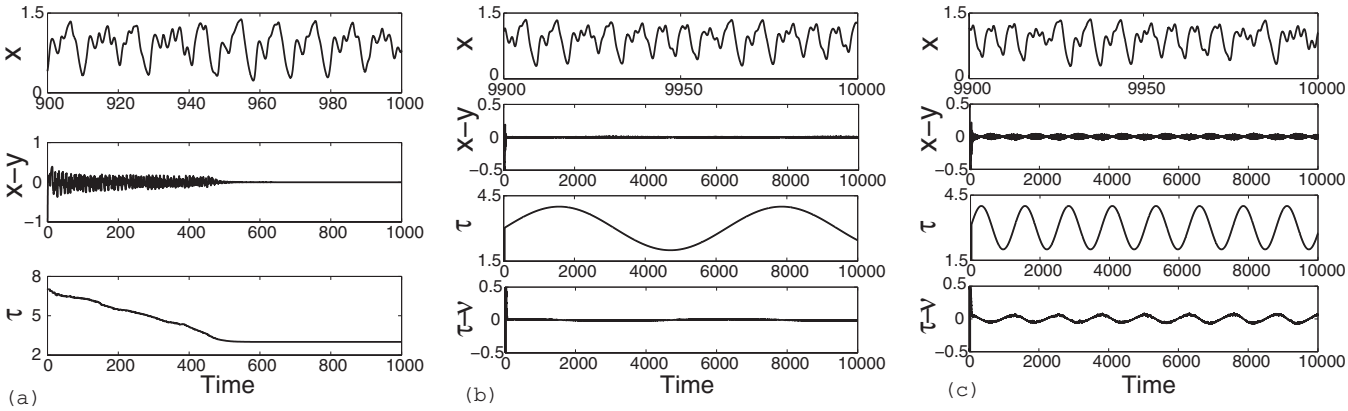

FIG. 1. Numerical results for Example 1 with the parameters chosen as $r=1$, $\delta=0.5$, and the initial values settled as $x(\theta)=0.1$, $y(\theta)=2$, $\epsilon(\theta)=0$, for all $\theta\in[-7,0]$. (a) The chaotic dynamics of model (7), the complete synchronization between signal $x(t)$ and $y(t)$, and the convergence of $\nu(t)$ showing an accurate identification of $\tau=3$. (b) A successful identification of the time-varying delay when the fluctuating frequency w of the time delay $\tau(t)$ is lower. (c) The estimator $\nu(t)$ oscillates around the true value of the time delay $\tau(t)$, showing an unsuccessful identification when the fluctuating frequency w is higher.

cal model for producing higher-dimensional chaos. Without the time delay, this one-dimensional model can never generate chaos. With the parameters $a=2$, $b=10$, $c=1$, and $\tau=3$, model (7) exhibits chaotic dynamics, as shown in Fig. 1(a). In the light of the proposed adaptive strategy, the corresponding coupling model and adaptive rules are designed as

$$\dot{y} = \frac{2y(t-\nu)}{1+[y(t-\nu)]^{10}} - y + \epsilon(y-x),$$

$$\dot{\nu} = -\delta(y-x)\partial_\nu F(y,\nu),$$

$$\dot{\epsilon} = -r(y-x)^2, \tag{8}$$

where

$$\partial_\nu F(y,\nu) = -\frac{2\{1+(1-10)[y(t-\nu)]^{10}\}}{\{1+[y(t-\nu)]^{10}\}^2}\left\{\frac{2y(t-2\nu)}{1+[y(t-2\nu)]^{10}} - y(t-\nu) + \epsilon(t-\nu)[y(t-\nu)-x(t-\nu)]\right\}.$$

As shown in Fig. 1(a), a complete synchronization between x and y is achieved with the parameters $r=1$ and $\delta=0.5$, and the estimator $\nu(t)$ converges to the exact values of $\tau=3$.

Furthermore, since time-varying delays are ubiquitous in real systems, online identification of time delays of this kind becomes a more challengeable problem. In order to test whether or not the proposed adaptive strategy is suitable for identifying the time-varying delays, we set τ in model (7) particularly as $\tau(t)=3+\sin(wt)$, where w is a parameter determining the fluctuating frequency of the delay. With this setting, model (7) still produces chaotic trajectories, as numerically shown in Figs. 1(b) and 1(c). When $w=10^{-3}$, that is, $\tau(t)$ fluctuates slowly, both the complete synchronization and the identification of the time-varying delays can be achieved, as shown in Fig. 1(b). Nonetheless, when $w=5\times10^{-3}$, that is, the fluctuating frequency is much higher, neither the complete synchronization nor the identification of time delay is achieved. As shown in Fig. 1(c), the estimator $\nu(t)$ never converges to $\tau(t)$ but oscillates slightly around the certain value $\tau=3$.

These numerical results can be explained as follows. The dynamics of the coupled models with the adaptive rules, though being convergent, have different convergence rate for different set of parameters and initial values. When the fluctuating frequency of $\tau(t)$ is below the convergence rate of the whole dynamics, there is an adequate duration for the estimator to trace the time-varying delay, analogous to the case of the time-invariant delay. Inversely, a much higher fluctuating frequency of $\tau(t)$ shortens the tracing duration, which results into the failure of both the complete synchronization and the identification. Therefore, the proposed adaptive rule can be used to identify the time-varying delays with a relatively lower fluctuating frequency. In addition, although those time-varying delays with very fast frequencies cannot be identified exactly through the proposed adaptive strategy, tuning those parameters such as r and m in Eq. (8) can somewhat accelerate the convergence rate of the whole dynamics and thus increase the tolerance value of the fluctuating frequency for convergence. We have a conjecture that the tolerance value for convergence is sensitively connected to the configuration of a time-delayed dynamical model, which deserves a further investigation.

Example 2. A time-delayed model with periodic dynamics. To illustrate the technical definition, "identifiability," proposed in Sec. II and explicitly discussed in Appendix B, consider a single neuron model with a discrete time delay [21], which is described by

$$\dot{x} = -x + h[ax - bx(t-\tau) + P]. \tag{9}$$

Here, the neuron activation function $h[u]$ is settled as

$$h[u] = \frac{1}{1+e^{-u}}. \tag{10}$$

When the parameters are chosen as $a=4$, $b=4.8$, $P=-0.8$, and $\tau=3$, model (9) produces periodic dynamics with an approximate period $T\approx9.46$, as shown in Fig. 2(a). Analo-

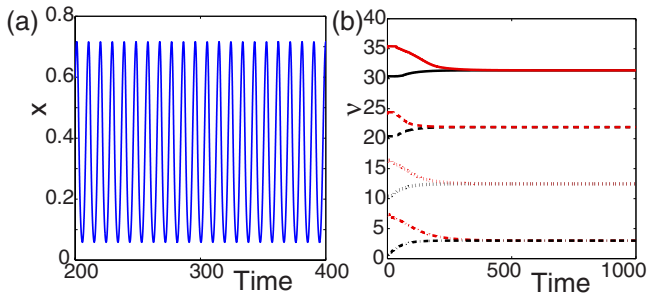

FIG. 2. (Color online) Numerical results for Example 2 in which the parameters are chosen as $r=5$, $\delta=5$, and the initial values are set as $x(\theta)\equiv 1$, $y(\theta)\equiv 2$, and $\epsilon(\theta)\equiv 0$ for all $\theta\in[-40,0]$. (a) A periodic trajectory of model (9) with an approximate period $T\approx 9.46$. (b) Dynamics of the estimator $\nu(t)$ starting from different initial values, $\nu(\theta)\equiv\nu_0^q(\theta\in[-40,0], q=1,\ldots,8)$. Here, ν_0^q are randomly chosen as 0, 7, 10, 16, 20, 24, 30, and 35, respectively. The convergence values of $\nu(t)$ correspondingly are 3, 3, 12.46, 12.46, 21.93, 21.93, 31.93, and 31.39, which all satisfy the form $\tau+n_0T$ for some nonnegative integer n_0.

gously, the coupling model with the adaptive rules is designed as

$$\dot{y} = -y + h[ay - by(t-\nu) + P] + \epsilon(y-x),$$

$$\dot{\nu} = -\delta(y-x)\partial_\nu F(y,\nu),$$

$$\dot{\epsilon} = -r(y-x)^2, \tag{11}$$

where

$$\partial_\nu F(y,\nu) = \frac{be^{-ay+by(t-\nu)-P}}{[1+e^{-ay+by(t-\nu)-P}]^2}\{-y(t-\nu) + h[ay(t-\nu) - by(t-2\nu) + P] + \epsilon(t-\nu)[y(t-\nu) - x(t-\nu)]\}.$$

As interestingly shown in Fig. 2(b), with randomly selected initial values, the estimator $\nu(t)$ always converges to a value equal to the true value of the time delay ($\tau=3$) plus some number that is nonnegative integer multiple of the period T. Clearly, this numerical result is consistent completely with one of the situations discussed in Appendix B. As indicated in Appendix B, when the observable synchronized manifold consists of periodic orbits, the vector field of the model sometimes is not exactly identifiable along the synchronized manifold. An accurate identification of the time delays in such systems cannot be always achieved unless the selected initial value for the estimator is relatively closed to the real value of the time delay. However, identification of this kind can be seen as a successful realization in the generalized sense, that is, the estimator $\nu(t)$ converges to $\tau+n_0T$ for some nonnegative integer n_0.

Example 3. A two-dimensional time-delayed model with multiple time delays and parameters. In real application, instead of estimating either the time delay or the parameter individually, one often needs to identify them simultaneously. In particular, with the increase of the number of the unknown time delays and parameters in a model, an accurate identification of them becomes more difficult and time-consuming. However, the above-proposed strategy allows us to identify them simultaneously and swiftly. To illustrate this, we consider the following neural network model with multiple time delays and parameters [22]:

$$\dot{x}_1(t) = -x_1 + \sum_{i=1}^{2} a_{1i}h[x_i(t)] + \sum_{i=1}^{2} b_{1i}h[x_i(t-\tau_{1i})] + I_1,$$

$$\dot{x}_2(t) = -x_2 + \sum_{i=1}^{2} a_{2i}h[x_i(t)] + \sum_{i=1}^{2} b_{2i}h[x_i(t-\tau_{2i})] + I_2, \tag{12}$$

where the neuron activation function $h[u]$ is the same as the function in Eq. (10), and the connection weight matrices are given as

$$\boldsymbol{A} = \{a_{ij}\}_{2\times 2} = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}, \quad \boldsymbol{B} = \{b_{ij}\}_{2\times 2} = \begin{bmatrix} -5 & -4 \\ -3 & -6 \end{bmatrix}.$$

In addition, the external current input $\boldsymbol{I}=[I_1,I_2]^{\mathrm{T}}$ is selected as a constant vector $[2,3]^{\mathrm{T}}$, and the signal transmission time delay matrix is set as

$$\boldsymbol{\tau} = \{\tau_{ij}\}_{2\times 2} = \begin{bmatrix} 15 & 2 \\ 12 & 5 \end{bmatrix}.$$

With these settings, two-dimensional model (12) generates a chaotic attractor as numerically displayed in Fig. 3(a).

Let a_{ij} and $\tau_{ij}(i,j=1,2)$ be the parameters and time delays pending for identification, respectively, and then design the coupling model with the adaptive rules as follows:

$$\dot{y}_1(t) = -y_1 + \sum_{i=1}^{2} c_{1i}h[y_i(t)] + \sum_{i=1}^{2} b_{1i}h[y_i(t-\tau_{1i})] + I_1 + \epsilon_1(y_1 - x_1),$$

$$\dot{y}_2(t) = -y_2 + \sum_{i=1}^{2} c_{2i}h[y_i(t)] + \sum_{i=1}^{2} b_{2i}h[y_i(t-\nu_{2i})] + I_2 + \epsilon_2(y_2 - x_2),$$

$$\dot{c}_{ij}(t) = -\gamma_{ij}(y_i - x_i)h[y_j(t)],$$

$$\dot{\nu}_{ij}(t) = -\delta_{ij}(y_i - x_i)\,\partial F_\nu(y_i,\nu_{ij}),$$

$$\dot{\epsilon}_i(t) = -r_i(y_i - x_i)^2, \tag{13}$$

where $i,j=1,2$, and

$$\partial F_\nu(y_i,\nu_{ij}) = \frac{b_{ij}e^{-y_i(t-\nu_{ij})}}{[1+e^{-y_i(t-\nu_{ij})}]^2}\dot{y}_i(t-\nu_{ij}). \tag{14}$$

Thus, Fig. 3(b) shows that the state variables $\boldsymbol{x}=[x_1,x_2]^{\mathrm{T}}$ and $\boldsymbol{y}=[y_1,y_2]^{\mathrm{T}}$ approach a complete synchronization. Figures 3(c) and 3(d) show that all the estimators $\nu_{ij}(t)$ and $c_{ij}(t)$ converge to the corresponding elements of $\boldsymbol{\tau}$ and $\boldsymbol{A}$ exactly and swiftly. Therefore, by the proposed strategy, the simultaneous identification of totally eight time delays and parameters in model (12) is achieved.

It is clear that the state variables are all utilized to design coupling model (13) in the above simulation. However, it is

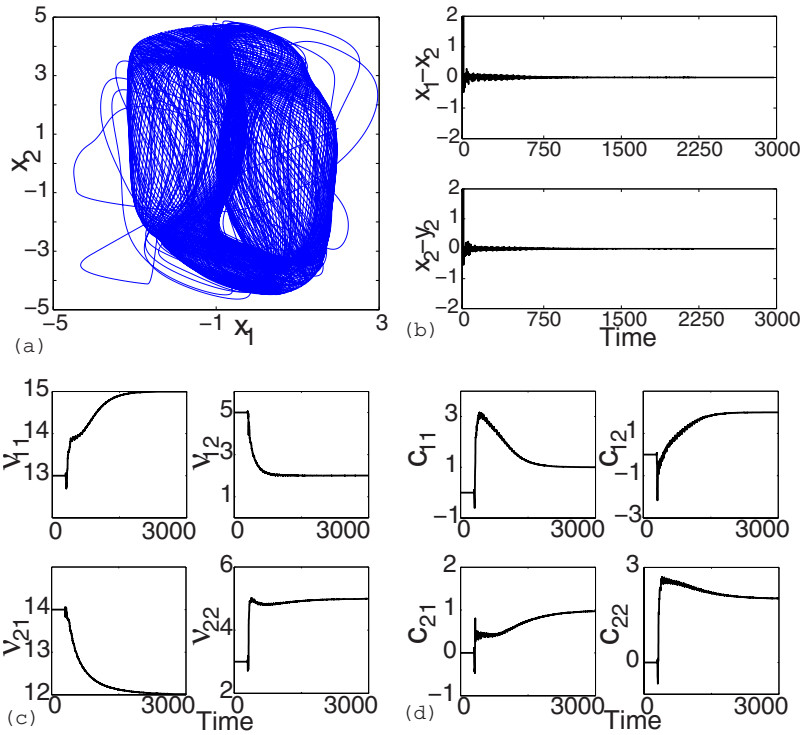

FIG. 3. (Color online) Numerical results for Example 3. (a) A chaotic attractor generated by model (12). (b) The error dynamics versus the time evolution with the parameters chosen as $r_{ij}=\gamma_{ij}=\delta_{ij}=1$, and with the initial values settled as $x_1(\theta)=2$, $x_2(\theta)=1.2$, $y_1(\theta)=-5$, $y_2(\theta)=-5$, $\epsilon_1(\theta)=0$, and $\epsilon_2(\theta)=0$ $(\theta\in[-15,0])$. [(c), (d)] The convergent dynamics of $\nu_{ij}(t)$ and $c_{ij}(t)(i,j=1,2)$ imply a successful identification of totally eight time delays and parameters, where the initial values for these estimators are randomly chosen.

neither necessary nor realistic that all state variables are taken into account in practice. As a matter of fact, in many physical, chemical and biological models, only part of the state variables can be detected. Thus, it is of practical significance to develop some feasible identification strategy only requiring the partial information from the state variables. To show that the proposed strategy is a strategy of this kind, rewrite the state variables vector $\boldsymbol{x}$ as $\boldsymbol{x}=[\boldsymbol{x}_1,\boldsymbol{x}_2]^{\mathrm{T}}$, and denote by $\boldsymbol{x}_{\mathrm{out}}=G(\boldsymbol{x}_2)$ the observed output signal of the model. Here, G is assumed as an observable function of variable $\boldsymbol{x}_2$, and the remainder $\boldsymbol{x}_1$ is supposed to be some internal signal that is not available or not of real interest. The problem becomes how to only use the output signal $\boldsymbol{x}_{\mathrm{out}}$ for realizing identification. For this specific example, we simply let x_2 in model (12) be the output signal, and let x_1 be the internal signal. Assume that τ_{11} and τ_{12} with all the parameters are known *a prior*, and that τ_{21} and τ_{22} are the unknown time delays pending for identification. Then the partially coupling model with the adaptive rules is designed as follows:

$$\dot{y}_1(t) = -y_1 + \sum_{i=1}^{2} a_{1i} h[y_i(t)] + \sum_{i=1}^{2} b_{1i} h[y_i(t-\tau_{1i})] + I_1,$$

$$\dot{y}_2(t) = -y_2 + \sum_{i=1}^{2} a_{2i} h[y_i(t)] + \sum_{i=1}^{2} b_{2i} h[y_i(t-\nu_{2i})] + I_2 + \epsilon(y_2 - x_2),$$

$$\dot{\nu}_{21}(t) = -\delta_1(y_2 - x_2)\,\partial F_\nu(y_2, \nu_{21}),$$

$$\dot{\nu}_{22}(t) = -\delta_2(y_2 - x_2)\,\partial F_\nu(y_2, \nu_{22}),$$

$$\dot{\epsilon}(t) = -r(y_2 - x_2)^2, \tag{15}$$

where ∂F_ν is the same function as defined in Eq. (14). Consequently, successful identification of both τ_{21} and τ_{22}, as well as a complete synchronization between coupled models (12) and (15), is realized numerically (see Fig. 4). It is worthwhile to mention that, only with partial information, an accurate identification of all time delays and parameters in the above model cannot be surely achieved by the proposed strategy. Further investigation is expected on what and how factors affect the time-delays identification with partial observables.

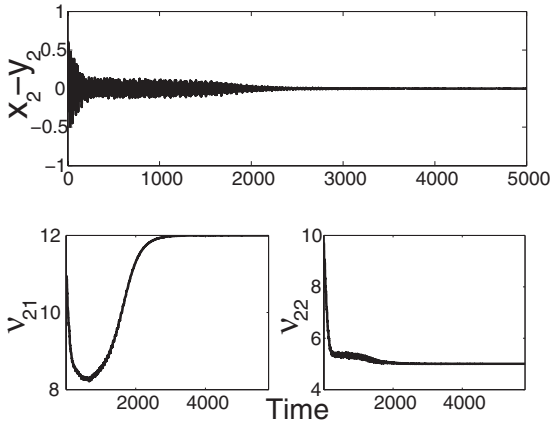

FIG. 4. A complete synchronization between coupled models (12) and (13), and a successful identification of time delays τ_{21} and τ_{22} only with a partial coupling configuration. Here, the parameters are chosen as $r=1$ and $\delta=1$. All the initial values are selected as $x_1(\theta)=2$, $x_2(\theta)=1.2$, $y_1(\theta)=y_2(\theta)=-5$, and $\epsilon(\theta)=0$ for all $\theta\in[-15,0]$.

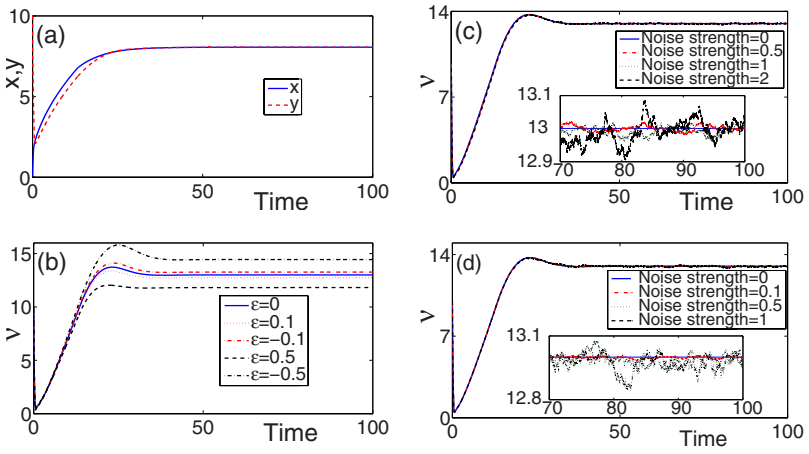

FIG. 5. (Color online) Successful complete synchronization between models (16) and (17) shown in (a). Accurate time-delay identification shown in (b) as the noise strength is zero. Here, the model's equilibrium is asymptotically stable and the time delay is in a distributed form. The dynamics of the estimator ν shown in (b), (c), and (d), respectively, for the parameter mismatch in model (17), for the additive noise in model (16), and for the multiplicative noise in model (16). All the cases illustrate the robustness of the proposed adaptive strategy against a certain amount of perturbations. Here, the parameters are taken as $r=1$ and $\delta=0.1$. All the initial values are selected as $x(\theta)=0.1$, $y(\theta)=10$, and $\epsilon(\theta)=0$ for all $\theta \in [-15,0]$.

Example 4. A time-delayed model with a distributed time delay. Here, we first validate the proposed strategy for identifying the distributed time delays in nonlinear models with stable dynamics. Let us consider a one-dimensional time-delayed system, which is described by

$$\dot{x} = -x^3 + \sigma \int_{-\tau}^{0} x(t+\xi)d\xi. \tag{16}$$

With $\sigma=5$ and $\tau=13$, it is easy to obtain that $x_t^*(\theta) \equiv \sqrt{65}$ is an equilibrium of model (16). In particular, numerical simulation in Fig. 5(a) manifests that this equilibrium is asymptotically stable.

According to the proposed strategy, the coupling model with the adaptive rules is constructed through

$$\dot{y} = -y^3 + \tilde{\sigma} \int_{-\nu}^{0} y(t+\xi)d\xi + \epsilon(y-x),$$

$$\dot{\nu} = -\delta(y-x)\tilde{\sigma}y(t-\nu),$$

$$\dot{\epsilon} = -r(y-x)^2, \tag{17}$$

where $\tilde{\sigma}$ is set to be 5. As depicted in Fig. 5(b), the distributed time delay can be successfully identified in spite of the existence of the asymptotically stable equilibrium $x_t^*(\theta)$. Interestingly, this numerical result is in agreement with the theoretical results concluded in Appendix B. It is further noted that, though the time delay in model (16) appears simply in a linear form, it is still impractical to implement the methods proposed in [9,10] for the time-delay identification. The reason is that the infinite and location-varying peaks, which are emergent in the plots of the correlationlike function, most likely lead the counting method to a failure, when the distributed time delays instead of the discrete time delays are taken into account.

Finally, we test the robustness of the proposed strategy for identifying the distributed time delay in several cases. Consider a case of parameter mismatch, that is, $\tilde{\sigma}$ in Eq. (17) is set to be $5+\varepsilon$, where ε is regarded as a mismatch level. As shown in Fig. 5(b), the accuracy of the identified time delay depends on the mismatch level. The smaller the level, the more accurate the identified time delay. Additionally, consider two types of noise perturbations in model (16): one is the additive noise, that is, the term $\rho\dot{W}$ is added directly to the left-hand side of Eq. (16); the other is the multiplicative noise, that is, σ in Eq. (16) is set as $5+\rho\dot{W}$. Here, W is a one-dimensional Brownian motion, $\dot{W}$ is a white noise accordingly, and ρ is a noise strength. Hence, the data produced by model (16) with noise perturbations are imported into coupling model (17) for time-delays identification. As clearly shown in the insets of Figs. 5(c) and 5(d), the inputs of the two types of noise perturbations induce the fluctuations of the estimator ν; however, the fluctuations eventually evolve in the vicinity of the true value $\tau=13$, when the noise strengths are not very strong. Further numerical simulations, which are not included here for conciseness, manifest that the data reproduced by the model with the identified time delays in the above cases can qualitatively reflect the dynamics of the data that are imported into the coupling model.

IV. TIME-DELAY IDENTIFICATION IN A GENE TRANSCRIPTION MODEL

The objective of this section is to utilize the above-developed adaptive synchronization strategy with biological data to identify the time delay in a gene transcription model and then to reproduce the dynamics that are qualitatively consistent with the experimental data.

Nowadays, it has been generally accepted that time delays resulting from transcription, transcript splicing and processing, and protein synthesis are omnipresent in the duration of mRNA and protein expression, and that, aside from additional components in feedback loop, time delay can result in various oscillations related to the circadian clocks of gene expression level [23–26]. Here, we consider a dynamical model, established in [25,26], describing the two-hour oscillation of mRNAs for Notch signaling molecules that are important for coordinated somite segmentation. The experimentally recorded time series on the oscillation of *hes1* mRNA in cultured cells, depicted in Fig. 6(a), are first reported in [25]. The corresponding dynamical model, based on the elementary knowledge of gene expression level, is given by

$$\dot{M} = \alpha_m G[P(t-\tau)] - \mu_m M,$$

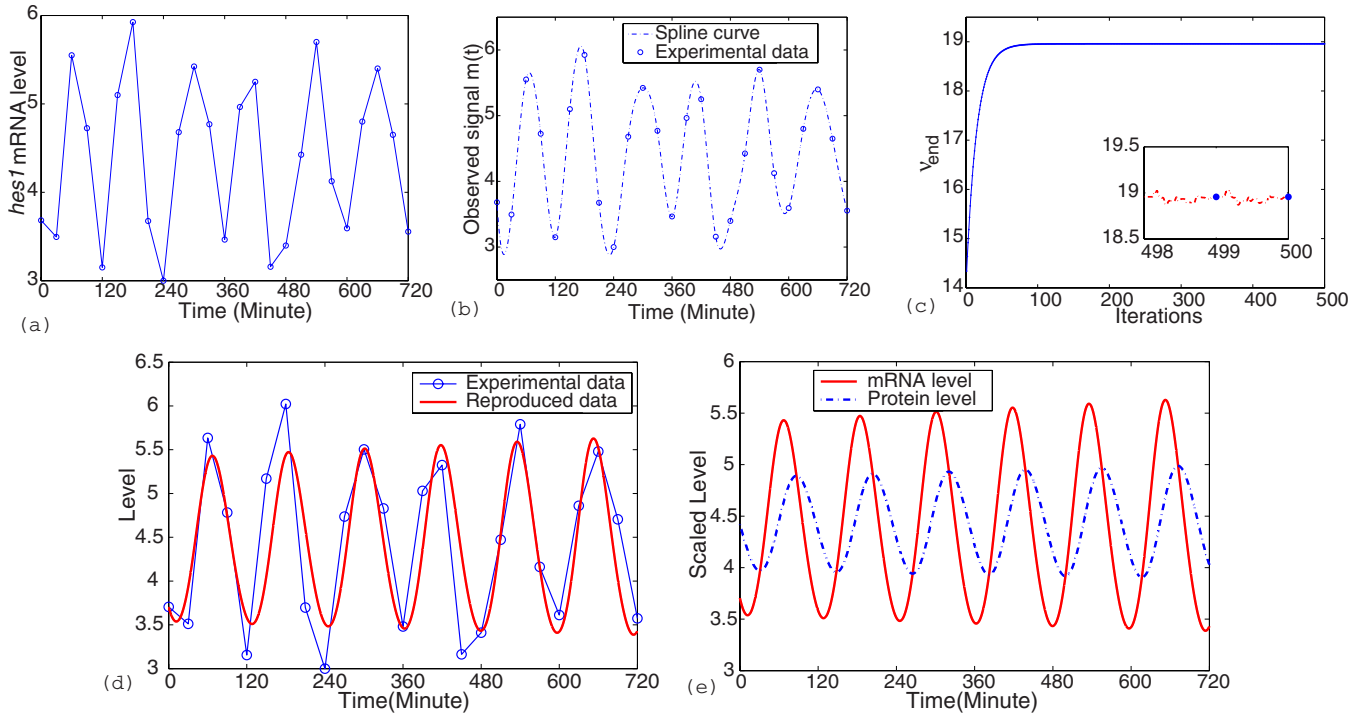

FIG. 6. (Color online) (a) The experimental data on the oscillating concentrations of *hes1* mRNA for Notch signaling molecules with almost two-hour cycles [25]. (b) Using the spline fit with the experimental data in (a) yields a completed time series which is presumably generated by $m(t)$. (c) Convergent dynamics of the end of the estimator ν in each recursive iteration, where ν_{end} converges gradually to 18.95. Inset: the estimator ν oscillates stably and slightly around 18.95 in the last two iterations, where the duration of each iteration is 720 min. [(d), (e)] Model (19), with the estimated time delay τ=18.95, generates oscillations of *hes1* mRNA and Hes1 protein, which coincide qualitatively with the experimental results reported in [25,26].

$$\dot{P} = \alpha_P M - \mu_P P, \tag{18}$$

where M and P denote, respectively, the concentrations at time t of *hes1* mRNA and Hes1 protein, τ represents the sum of the transcriptional and translational time delays, and G is a monotonic decreasing function, representing the delayed repression of *hes1* mRNA production by Hes1 protein and taking the form

$$G[P(t-\tau)] = \frac{1}{1+[P(t-\tau)/P_0]^n}.$$

Here, P_0 is the repression threshold and n is the Hill coefficient. In addition, α_m, α_P, μ_m, and μ_P are the rate constants of biological relevance. Now, with a set of new variables,

$$m = M/\alpha_m, \quad p = P/\alpha_m\alpha_P, \quad p_0 = P_0/(\alpha_m\alpha_P),$$

rescaling model (18) yields

$$\dot{m} = g[p(t-\tau)] - \mu_m m,$$

$$\dot{p} = m - \mu_P p, \tag{19}$$

where $g[p(t-\tau)]=\frac{1}{1+[p(t-\tau)/p_0]^n}$. As illustrated in [26], this rescaled model brings much convenience in the analysis of the model dynamics with the parameters and time delay. Analogous to [26], the parameters are, *a prior*, settled as $\mu_m=\mu_P=0.03/\text{min}$, $p_0=100$, and $n=5$, and the time delay τ around 15–20 min is left for identification.

We use the adaptive strategy to identify the time delay in model (19). To this end, the partially coupling model with the adaptive rules is designed as follows:

$$\dot{y}_1 = g[y_2(t-\nu)] - \mu_m y_1 + \epsilon(y_1 - m),$$

$$\dot{y}_2 = y_1 - \mu_P y_2,$$

$$\dot{\nu} = -\delta(y_1 - m)M(t),$$

$$\dot{\epsilon} = -r(y_1 - m)^2, \tag{20}$$

where

$$M(t) = \frac{[y_2(t-\nu)/100]^4[y_1(t-\nu) - \mu_P y_2(t-\nu)]}{20\{1+[y_2(t-\nu)/100]^5\}^2},$$

and the driving signal $m(t)$ represents the time series obtained in the experiments. However, as shown in Fig. 6(a), the sampling frequency of the experimental time series is too low to directly adopt the adaptive strategy. For this reason, we first generate a spline curve passing through all the points of the experimental time series, and then use the points from t=0 to 720 min on this curve to construct a completed time series that is presumably generated by $m(t)$ [see Fig. 6(b)]. Also to make a good use of this completed time series, we recursively adopt the adaptive strategy in Eq. (20) such that the end value of the estimator ν in each recursive iteration is set as the initial value for the next iteration. Here, the dura-

tion of each iteration is 720 min. As shown in Fig. 6(c), the end value of the estimator ν approaches a constant 18.95 after 500 iterations. It is noted that the estimator ν still fluctuates during each iteration. Nevertheless, with the increase of the iteration times, the estimator ν shows a stable oscillation. In particular, as shown in the inset of Fig. 6(c), the oscillation of the estimator ν, in the last two iteration times, fluctuates in the vicinity of 18.95, which is approximately equal to the average number of ν in these iteration times. Now, with 18.95 which is regarded as an approximate estimation of the time delay τ, we depict the dynamics of both *hes1* mRNA and Hes1 protein in Figs. 6(d) and 6(e). The dynamics produced coincide qualitatively with the experimental data in [25]. However, no oscillation is in agreement with the experimental data when the time delay is either excluded or far from the identified value, as indicated in [26]. Consequently, appropriate use of the adaptive synchronization strategy with the truly experimental data allows us to not only identify the time delay approximately but also capture the dynamics of real systems qualitatively.

We note that the estimator does not converge perfectly during the last two iterations, and the reproduced data for *hes1* mRNA protein are not exactly consistent with the experimental data, as shown in the inset of Figs. 6(c) and 6(d), respectively. These imperfections in estimators convergence and data reproduction are unavoidable, since various types of noise perturbations intrinsically exist in the gene transcription and experimental observations. However, the qualitative consistency shown above results from the adaptive strategy developed, when used with data from real systems, is robust against a certain amount of perturbation.

V. CONCLUDING REMARKS

This paper proposes an adaptive synchronization strategy for identifying various types of time delays in nonlinear dynamical models. The effectiveness of this adaptive strategy is theoretically validated by the theory of functional differential equations. More significantly, the problem of the identifiability of a given vector field is introduced and then extensively investigated for diverse forms of time delays. Several time-delayed models, as well as the gene transcription model with real data, are used to illustrate the practical usefulness and the restrictions of the proposed strategy. It is clear that the computational complexity is only $O(L)$ if the time delays can be adaptively identified. We therefore expect that the proposed strategy for time delays identification as well as the existing techniques for parameters estimation could be widely used in practice. Moreover, it is stressed here that our proposed strategy relies on the specific forms of the models, where the parameters and time delays are allowed are totally unknown. Modification methods need further development for the models that, with identified parameters and time delays, comes into conflict with the experimental evidence qualitatively.

The results in the paper also suggest the following directions for future research. From a theoretical direction, a thorough investigation on the proposed strategy for a more general class of time-delayed dynamical models would be useful. Analytical criteria need to be further established for time-delay identification in models with either multiple discrete time delays (say $p_i > 2$, as mentioned in Appendix B) or with more general forms of distributed time delays. From experimental and numerical directions, time-delay identification and parameter estimation should jointly work with the noise perturbed and high throughput data that are available in physical, chemical, and biological studies. Also, tune-up methods of the parameters in the proposed strategy, such as δ_{ij} and r_i, need developing for a practical identification. Lastly, identification of the models whose forms are completely unknown is of great importance, and deserves future investigation.

ACKNOWLEDGMENTS

W.L. was supported by LMNS and the NNSF of China (Grants No. 60874121 and No. 10971035), and partially by the 973 Program of China (Grant No. 2007CB814904). B.X. was supported by the NNSF of China (Grant No. J0730103). J.F. was supported by an EPSRC (U.K.) grant CARMAN and an EU grant BION. Authors are grateful to Professor J. Socola and Professor D. Waxman for their very significant suggestions on this work and providing important references leading to many improvements.

APPENDIX A: A COMPACT FORM OF THE VECTOR FIELD

Mathematically, the vector field in Eq. (2) can be rewritten in the following uniform and compact form:

$$F_i(\boldsymbol{\varphi}) = \sum_{j=1}^{p_i} \int_{-\tau_M}^{0} \rho_{ij}(\theta) f_{ij}(\boldsymbol{\varphi}(\theta)) d\theta, \tag{A1}$$

where each p_i is a finite natural number, $\boldsymbol{\varphi} \in \mathbf{C}([-\tau_M, 0], \mathbb{R}^n)$, and each $\rho_{ij}(\theta)$ belongs to the function family $\mathcal{F}([-\tau_M, 0], \mathbb{R}^+) = \mathcal{D}([-\tau_M, 0], \mathbb{R}^+) \cup \mathcal{X}([-\tau_M, 0], \mathbb{R}^+)$. Here, $\mathcal{D}([-\tau_M, 0], \mathbb{R}^+)$ consists of all the Dirac delta functions with a point-set support in the interval $[-\tau_M, 0]$, and $\mathcal{X}([-\tau_M, 0], \mathbb{R}^+)$ consists of all the eigenfunctions satisfying $\rho_{ij}(\theta) = 1$ when θ belongs to the interval $[-\tau_{ij}, 0]$ $(0 \le \tau_{ij} \le \tau_M)$ and $\rho_{ij}(\theta) = 0$ when θ belongs to $[-\tau_M, -\tau_{ij})$. Specifically, when $\rho_{ij}(\theta)$ is selected as the Dirac delta function with a point-set support $\{-\tau_{ij}\}$, in which $\tau_{ij} \in (0, \tau_M]$, model (1) has multiple discrete time delays, that is, $\int_{-\tau_M}^{0} \rho_{ij}(\theta) f_{ij}(\boldsymbol{\varphi}(\theta)) d\theta = f_{ij}(\boldsymbol{\varphi}(-\tau_{ij}))$. If all $\tau_{ij} = 0$, it becomes a trivial case that model (1) has no time delays. In addition, when $\rho_{\hat{i}\hat{j}}(\theta)$ belongs to $\mathcal{X}([-\tau_M, 0], \mathbb{R}^+)$ for some indexes $\hat{i}$ and $\hat{j}$, model (1) has at least one time delay in the distributed form, namely, $\int_{-\tau_M}^{0} \rho_{\hat{i}\hat{j}}(\theta) f_{\hat{i}\hat{j}}(\boldsymbol{\varphi}(\theta)) d\theta = \int_{-\tau_{\hat{i}\hat{j}}}^{0} f_{\hat{i}\hat{j}}(\boldsymbol{\varphi}(\theta)) d\theta$.

APPENDIX B: IDENTIFIABILITY OF THE VECTOR FIELD

The identifiability of a given vector field relies crucially on the specific form of its components $F_i (i = 1, \ldots, n)$ as well as on the structure of the observable synchronized manifold. However, several feasible criteria of physical significance can still be established.

For simplicity but without loss of generality, let $p_i=2$ and $\rho_{i1} \in \mathcal{D}([-\tau_M,0],\mathbb{R}^+)$ with a support $\{0\}$ in Eq. (A1). Then, we have

$$F_i(\boldsymbol{x}_t) = f_{i1}(\boldsymbol{x}(t)) + \int_{-\tau_M}^{0} \rho_i(\theta) f_{i2}(\boldsymbol{x}_t) d\theta,$$

where the first term on the right hand side contains no time delay. On the one hand, for $\rho_{i2} \in \mathcal{D}([-\tau_M,0],\mathbb{R}^+)$, we have coupled models with discrete time delays. If each function f_{i2} is globally invertible, then Eq. (6) in Sec. II implies

$$\boldsymbol{x}(t-\xi_{i2}^*) = \boldsymbol{x}(t-\tau_{i2}), \tag{B1}$$

for all t and all $\boldsymbol{x}_t(\theta)$ on the observable synchronized manifold of model (1). Notice that the signals of interest on the observable synchronized manifold are bounded in general. Hence, the following three cases for time-delay's identification are taken into account:

(1) If the observable synchronized manifold only contains the equilibrium of model (1), i.e., $\boldsymbol{x}_t(\theta) \equiv \boldsymbol{x}^*$, then ξ_{i2}^* is equal to any number including the value of τ_{i2}. This means that the vector field is not identifiable, and that time-delay's identification becomes failed almost surely for this case.

(2) If the observable synchronized manifold consists of periodic orbit with period T, i.e., $\boldsymbol{x}(t)=\boldsymbol{x}(t+T)$ for all t, then Eq. (B1) implies $\xi_{i2}^* \in \{\tau_{i2}+nT \in [0,\tau_M] | n \in \mathbb{N} \cup \{0\}\}$. Notice that $\{\tau_{i2}+nT \in [0,\tau_M] | n \in \mathbb{N} \cup \{0\}\}$ is a discrete point set, and that ω-limit set is connected. Hence, each estimator $\nu_{i2}(t)$ can only converge to $\tau_{i2}+n_0T$ for some nonnegative integer n_0, where the value of n_0 depends on the choice of all the initial values for the dynamical model. This can be regarded as a situation where the generalized identification of the time delay is realized, so that the vector field is identifiable in the generalized sense.

(3) If the observable synchronized manifold is chaotic, then Eq. (B1) implies $\xi_{i2}^*=\tau_i$. This manifests that the vector field is identifiable, and that the time-delay's identification can be surely achieved for this case.

If the function f_{i2} is not globally invertible, the above argument can still be used in the local neighborhood of the existence of its inverse.

On the other hand, for $\rho_{i2} \in \mathcal{X}([-\tau_M,0],\mathbb{R}^+)$, we have coupled models with distributed time delays. Equation (6) implies

$$C(t) \triangleq \int_{-\tau_{i2}}^{-\xi_{i2}^*} f_{i2}(\boldsymbol{x}_t(\theta)) d\theta = \int_{t-\tau_{i2}}^{t-\xi_{i2}^*} f_{i2}(\boldsymbol{x}(s)) ds = 0, \tag{B2}$$

for all t and all $\boldsymbol{x}_t(\theta)$ on the observable synchronized manifold. Then, two situations are considered as follows.

(1) If this manifold only contains an equilibrium $\boldsymbol{x}^*$ and each f_{i2} is nonzero at this equilibrium, we have $\xi_{i2}^*=\tau_{i2}$ for all i, so that the vector field is identifiable. Interestingly, this is different from the above case that discrete time-delayed models only with stable equilibria are not identifiable.

(2) If the manifold contains either periodic or chaotic attractors and f_{i2} is globally invertible, then the problem on the identifiability for the distributed time-delayed models can be transformed to the problem for discrete time-delayed models. To show this, we compute the derivative of $C(t)$ in Eq. (B2), which yields $\boldsymbol{x}(t-\xi_{i2}^*)-\boldsymbol{x}(t-\tau_{i2}) \equiv 0$. Clearly, the obtained equation is the same as Eq. (B1). Thus, for the periodic synchronized manifold, and with the condition $\int_0^T f_{i2}(\boldsymbol{x}(s))ds=0$ for periodic orbit $\boldsymbol{x}(t)$, the vector field is identifiable in the generalized sense as mentioned before; however, the vector field is identifiable if the integral condition is not valid. In addition, for the chaotic synchronized manifold, the vector field is identifiable.

We can further investigate the case of $p_i>2$ with more time delays, which however requires more intricate discussions, in that some additional assumptions, analogous to the linear independence condition proposed in [18], need importing. Numerical examples in Sec. III show successful identifications for such a case.

APPENDIX C: BOUNDEDNESS OF THE TRAJECTORY

Here, we are to investigate the boundedness of the trajectory generated by a particular kind of coupled models (1) and (3). For simplicity, take $\rho_i \in \mathcal{D}([-\tau_M,0],\mathbb{R}^+)$. Then, we obtain coupled models with discrete time delays τ_{ij}. Consider the following function:

$$V(t) = \sum_i \frac{1}{2}[y_i(t)-x_i(t)]^2 + \sum_{i,j} \frac{1}{2} \int_{t-\tau_{ij}}^{t} [f_{ij}(\boldsymbol{y}(s)) - f_{ij}(\boldsymbol{x}(s))]^2 ds + \sum_i \frac{1}{2r_i}[\epsilon_i(t)+L]^2 + \sum_{i,j} \frac{1}{2\delta_{ij}}[\nu_{ij}(t) - \tau_{ij}]^2,$$

for any $t \in [-\tau_M, t_0]$ and some $t_0 < +\infty$. Here, L is a positive constant pending for determination. Then, the derivative of $V(t)$ with respect to t along with the coupled models (1) and (3) yields

$$\begin{aligned}
\dot{V}(t) &= \sum_i [y_i(t)-x_i(t)] \Bigg\{ \sum_j [f_{ij}(\boldsymbol{y}(t-\nu_{ij})) - f_{ij}(\boldsymbol{x}(t-\tau_{ij}))] + \epsilon_i(t)[y_i(t)-x_i(t)] \Bigg\} + \sum_{i,j} \frac{1}{2} \{ [f_{ij}(\boldsymbol{y}(t)) - f_{ij}(\boldsymbol{x}(t))]^2 \\
&\quad - [f_{ij}(\boldsymbol{y}(t-\tau_{ij})) - f_{ij}(\boldsymbol{x}(t-\tau_{ij}))]^2 \} - \sum_i [\epsilon_i(t)+L][y_i(t) - x_i(t)]^2 - \sum_{i,j} [\nu_{ij}(t)-\tau_{ij}][y_i(t)-x_i(t)] \frac{\partial f_{ij}(\boldsymbol{y}(t-\upsilon_{ij}))}{\partial \upsilon_{ij}} \\
&= \sum_{i,j} [y_i(t)-x_i(t)][f_{ij}(\boldsymbol{y}(t-\nu_{ij})) - f_{ij}(\boldsymbol{y}(t-\tau_{ij}))] + \sum_{i,j} [y_i(t)-x_i(t)][f_{ij}(\boldsymbol{y}(t-\tau_{ij})) - f_{ij}(\boldsymbol{x}(t-\tau_{ij}))] \\
&\quad + \sum_{i,j} \frac{1}{2} \{ [f_{ij}(\boldsymbol{y}(t)) - f_{ij}(\boldsymbol{x}(t))]^2 - [f_{ij}(\boldsymbol{y}(t-\tau_{ij})) - f_{ij}(\boldsymbol{x}(t-\tau_{ij}))]^2 \} - \sum_i L[y_i(t)-x_i(t)]^2 - \sum_{i,j} [\nu_{ij}(t)-\tau_{ij}][y_i(t) \\
&\quad - x_i(t)] \frac{\partial f_{ij}(\boldsymbol{y}(t-\nu_{ij}))}{\partial \nu_{ij}}.
\end{aligned}$$

Now applying the elementary inequality $ab \leq \frac{1}{2}(a^2+b^2)$ to the second term after the above second equality gives

$$\dot{V}(t) \leq \sum_i \left(\frac{p_i}{2} - L\right)[y_i(t) - x_i(t)]^2 + \sum_{i,j} \frac{1}{2}[f_{ij}(\mathbf{y}(t)) - f_{ij}(\mathbf{x}(t))]^2 + \sum_{i,j} [y_i(t) - x_i(t)][f_{ij}(\mathbf{y}(t-\nu_{ij})) - f_{ij}(\mathbf{y}(t-\tau_{ij}))] - \sum_{i,j} [\nu_{ij}(t) - \tau_{ij}][y_i(t) - x_i(t)] \frac{\partial f_{ij}(\mathbf{y}(t-\nu_{ij}))}{\partial \nu_{ij}}.$$

Assume that the functions f_{ij} are globally Lipschitz with Lipschitz constants $l_{ij} > 0$, and that

$$f_{ij}(\mathbf{y}(t-\nu_{ij})) - f_{ij}(\mathbf{y}(t-\tau_{ij})) = \frac{\partial f_{ij}(\mathbf{y}(t-\nu_{ij}))}{\partial \nu_{ij}}(\nu_{ij} - \tau_{ij}) + O[(\nu_{ij} - \tau_{ij})^2],$$

where $O[\cdot]$ is a bounded quantity in the interval $[-\tau_M, t_0]$. With these assumptions, $\dot{V}(t)$ can be further estimated as follows:

$$\dot{V}(t) \leq \sum_i \left(-L + \frac{p_i}{2} + \sum_j l_{ij}\right)[y_i(t) - x_i(t)]^2 - \sum_{i,j} O[(\nu_{ij} - \tau_{ij})^2][y_i(t) - x_i(t)].$$

Thus, $\dot{V}(t) \leq 0$ is valid provided that L is sufficiently large and that the initial values of the coupled models satisfies $V(t) \leq \epsilon$ for $t \in [-\tau_M, 0]$, where ϵ is sufficiently small. Mathematically, we have the boundedness of the trajectory generated by the coupled models (1) and (3).

[1] J. Stark and K. Hardy, Science **301**, 1192 (2003).

[2] *Frontiers in Computational and Systems Biology*, edited by J. Feng, W. Fu, and F. Sun (Springer-Verlag, London, 2010); O. Wolkenhauer, in *Systems Biology: Dynamic Pathway Modelling*, edited by O. Wolkenhauer, P. Wellstead, and K. Ho. Cho (Portland Press, London, 2008); J. Keener and J. Sneyd, *Mathematical Physiology*, 2nd ed. (Springer, New York, 2009).

[3] A. L. Hodgkin and A. F. Huxley, Bull. Math. Biol. **52**, 25 (1990); C. Rocsoreanu, A. Georgescu, and N. Giurgiteanu, *The FitzHugh-Nagumo model: Bifurcation and Dynamics* (Kluwer Academic Publishers, Dordrecht, 2000); J. Rinzel, D. Terman, X.-J. Wang, and B. Ermentrout, Science **279**, 1351 (1998); A. Shpiro, R. Curtu, J. Rinzel, and N. Rubin, J. Neurophysiol. **97**, 462 (2007).

[4] S.-I. Niculescu, *Delay Effects on Stability* (Springer-Verlag, London, 2001).

[5] J. P. Richard, Automatica **39**, 1667 (2003).

[6] C. M. Marcus and R. M. Westervelt, Phys. Rev. A **39**, 347 (1989); U. Dressler and G. Nitsche, Phys. Rev. Lett. **68**, 1 (1992); K. Pyragas, Phys. Rev. E **58**, 3067 (1998); M. K. Stephen Yeung and S. H. Strogatz, Phys. Rev. Lett. **82**, 648 (1999); J. Ruan, L. Li, and W. Lin, Phys. Rev. E **63**, 051906 (2001); V. Pyragas and K. Pyragas, *ibid.* **73**, 036215 (2006); W. Just, B. Fiedler, M. Georgi, V. Flunkert, P. Hovel, and E. Scholl, *ibid.* **76**, 026210 (2007).

[7] J. Tuch, A. Feuer, and Z. Palmor, IEEE Trans. Autom. Control **39**, 823 (1994).

[8] S. V. Drakunov, W. Perruquetti, J. P. Richard, and L. Belkoura, Annu. Rev. Control **30**, 143 (2006).

[9] M. Siefert, Phys. Rev. E **76**, 026215 (2007).

[10] V. I. Ponomarenko and M. D. Prokhorov, Phys. Rev. E **78**, 066207 (2008); M. D. Prokhorov and V. I. Ponomarenko, *ibid.* **80**, 066206 (2009).

[11] M. J. Bünner *et al.*, Phys. Lett. A **211**, 345 (1996); M. J. Bünner, T. Meyer, A. Kittel, and J. Parisi, Phys. Rev. E **56**, 5083 (1997); H. Voss and J. Kurths, Phys. Lett. A **234**, 336 (1997); R. Hegger, M. J. Bünner, H. Kantz, and A. Giaquinta, Phys. Rev. Lett. **81**, 558 (1998); M. J. Bünner *et al.*, EPL **42**, 353 (1998); L. Cimponeriu, M. Rosenblum, and A. Pikovsky, Phys. Rev. E **70**, 046213 (2004).

[12] D. Yu and S. Boccaletti, Phys. Rev. E **80**, 036203 (2009); F. Sorrentino, *ibid.* **81**, 066218 (2010).

[13] J. Hale and S. Lunel, *Introduction to Functional Differential Equations* (Springer-Verlag, New York, 1993); J. Hale, *Theory of Functional Differential Equations* (Springer-Verlag, London, 2003).

[14] V. Kolmanovskii and A. Myshkis, *Applied Theory of Functional Differential Equations* (Kluwer Academic Publishers, Dordrecht, 1992).

[15] V. Kolmanovskii and A. Myshkis, *Introduction to the Theory and Applications of Functional Differential Equations* (Kluwer Academic Publishers, Dordrecht, 1999).

[16] U. Parlitz, Phys. Rev. Lett. **76**, 1232 (1996).

[17] A. Maybhate and R. E. Amritkar, Phys. Rev. E **61**, 6461 (2000); D. Huang, *ibid.* **69**, 067201 (2004); **71**, 037203 (2005); **73**, 066204 (2006); W. Yu, G. Chen, J. Cao, J. Lü, and U. Parlitz, *ibid.* **75**, 067201 (2007); W. Lin, H. Ma, J. Feng, and G. Chen, *ibid.* **82**, 046214 (2010).

[18] W. Lin and H.-F. Ma, Phys. Rev. E **75**, 066212 (2007); H. Ma and W. Lin, Phys. Lett. A **374**, 161 (2009).

[19] W. Lin, Phys. Lett. A **372**, 3195 (2008); W. Lin and H. Ma, IEEE Trans. Autom. Control **55**, 819 (2010).

[20] M. C. Mackey and L. Glass, Science **197**, 287 (1977).

[21] K. Gopalsamy and I. Leung, IEEE Trans. Neural Netw. **8**, 341 (1997).

[22] L. Olien and J. Belair, Physica D **102**, 349 (1997).

[23] P. Smolen, D. A. Baxter, and J. H. Byrne, Neuron **26**, 567 (2000).

[24] D. Bratsun, D. Volfson, L. S. Tsimring, and J. Hasty, Proc. Natl. Acad. Sci. U.S.A. **102**, 14593 (2005).

[25] H. Hirata *et al.*, Science **298**, 840 (2002).

[26] N. A. M. Monk, Curr. Biol. **13**, 1409 (2003).