Bifurcations and chaos in register transitions of excised larynx experiments

Isao T. Tokuda[a)]
School of Information Science, Japan Advanced Institute of Science and Technology, Ishikawa 923-1292, Japan

Jaromir Horáček
Institute of Thermomechanics, Academy of Sciences of the Czech Republic, Dolejškova 5, 182 00 Prague 8, Czech Republic

Jan G. Švec
Laboratory of Biophysics, Faculty of Science, Palacký University Olomouc, tř. Svobody 26, 771 46 Olomouc, Czech Republic

Hanspeter Herzel
Institute for Theoretical Biology, Humboldt University Berlin, Invalidenstr. 43, D-10115 Berlin, Germany

(Received 9 August 2007; accepted 21 November 2007; published online 14 January 2008)

Experimental data from an excised larynx are analyzed in the light of nonlinear dynamics. The excised larynx provides an experimental framework that enables artificial control and direct observation of the vocal fold vibrations. Of particular interest in this experiment is the coexistence of two distinct vibration patterns, which closely resemble chest and falsetto registers of the human voice. Abrupt transitions between the two registers are typically accompanied by irregular vibrations. Two approaches are presented for the modeling of the excised larynx experiment; one is the nonlinear predictive modeling of the experimental time series and the other is the biomechanical modeling (three-mass model) that takes into account basic mechanisms of the vocal fold vibrations. The two approaches show that the chest and falsetto vibrations correspond to two coexisting limit cycles, which jump to each other with a change in the bifurcation parameter. Irregular vibrations observed at the register jumps are due to chaos that exists near the two limit cycles. This provides an alternative mechanism to generate chaotic vibrations in excised larynx experiment, which is different from the conventionally known mechanisms such as strong asymmetry between the left and right vocal folds or excessively high subglottal pressure. © *2008 American Institute of Physics*. [DOI: 10.1063/1.2825295]

The vocal folds constitute a highly nonlinear self-oscillating system involving aerodynamic, biomechanical, physiological, and acoustic factors. The vocal registers, which are characterized by distinguished vibratory patterns of the vocal folds such as chest and falsetto, are controlled by a muscle activity as a bifurcation parameter of the nonlinear system. For a better understanding of the vocal registers, it is of significant importance to investigate the bifurcation structure of the vocal folds system. However, direct observation of such bifurcations in living human larynges is not simple, since it requires a high-speed camera inserted deeply into the throat in a highly invasive manner. In contrast, excised larynx experiments enable an artificial control of the vocal fold tension as well as a direct and detailed measurement of the vocal folds vibrations. Of particular interest in this experimental framework is the simulation of abrupt transitions between the chest and falsetto registers. Up to date, linear analysis such as the spectral analysis has been applied mainly to the excised larynx data, while only a limited analysis based on nonlinear dynamics has been carried out. The present paper applies two modeling approaches to the excised larynx experiment. One is a nonlinear modeling of measured time series and the other is a biomechanical modeling of the vocal folds vibrations. It is shown that the chest and falsetto vibrations are well characterized as limit cycle oscillations and their abrupt transitions are elucidated in terms of dynamical switching between the two coexisting limit cycles typically accompanied by chaos. This introduces a new concept of bifurcations to the register transitions in voice production.

I. INTRODUCTION

Voice production constitutes an important research area in science, engineering, medicine, and music. Complete modeling of the human voice is yet unsatisfactory, since the vocal fold oscillation is highly complex and nonlinear. According to the myoelastic-aerodynamic theory of voice production, the vocal fold oscillation is due to combined effects of subglottal pressure, airflow, elasticity of the vocal fold tissue, and collision between the vocal folds.[1] Fundamental frequency and glottal pulse shape are controlled by muscle action, which determines the effective length, mass, and tension of the vocal folds. The vocal registers, which define voiced sound quality, based on the vibratory pattern of the vocal folds such as chest and falsetto, are also due to the

[a)]Electronic mail: i.tokuda@biologie.hu-berlin.de.

strong control of the muscle activity. The glottal source signal is finally transformed into meaningful voiced speech through the vocal tract, which functions as a filter.

Considering the highly nonlinear oscillations of the vocal folds, the concept of nonlinear dynamics plays a key role for a better understanding of the human voice. The attractor types such as *limit cycle*, *subharmonics*, *biphonation*, and *chaos* provide a clear classification of the voiced sounds. Transitions between different types of the voice attractors can be well elucidated in terms of *bifurcations*. These ideas have been successfully applied to the studies of speech signal analysis,[2,3] vocal fold modeling,[4] voice pathology,[5,6] contemporary vocal music[7] and animal vocalizations.[8–12]

Nonlinear dynamics can be discussed not only with the real human voice but also with excised larynx experiments, which enable artificial control, direct observation, and detailed measurement of the vocal fold vibrations.[13,14] This experimental framework has been known to be of significant use for the study of laryngeal physiology and voice production. Berry *et al.*[15] controlled the asymmetry between the left and right vocal folds to study synchronous as well as nonsynchronous vibrations in the excised larynx experiment. Jiang and Zhang[16,17] applied an excessively high subglottal pressure to induce spatiotemporal chaos in the excised larynx. The main focus of the present paper is on another experimental framework that simulates transitions between chest and falsetto registers in the excised larynx.[18–20] This system produces two distinct vibratory patterns, which have good correspondence to the chest and falsetto registers of the human voice. Coexistence of the two vibrations induces abrupt transitions between them, sometimes accompanied by irregular vibrations as observed in real voice. Nonlinear analysis and modeling of such register transitions are of significant importance for the voice research, since dynamical mechanism of the register transitions have not yet been deeply understood. Hence, the aim of the present paper is to demonstrate that this excised larynx experiment can be well elucidated in terms of nonlinear dynamics. The register transitions can be considered as bifurcations of limit cycles that correspond to chest and falsetto vibrations and the irregular vibrations observed near the transitions are characterized as chaos. This presents a new mechanism to generate chaotic vibrations in excised larynx experiment, which is different from the conventionally known mechanisms such as strong asymmetry between the left and right vocal folds[6] or excessively high subglottal pressure.[21]

Two approaches are presented for the modeling of the excised larynx experiments. The first approach is based upon nonlinear predictive modeling of the time series data. It is shown that a relatively low-dimensional nonlinear model is capable of reproducing the essential features of the complex laryngeal vibrations including bifurcations of the chest and falsetto registers and chaos near the register transition.

The second approach is based upon biomechanical modeling of the excised larynx. We compare the biomechanical model that focuses on very basic mechanisms of the vocal fold vibration with the experiment. Such a simple model is shown to capture the essential features of the experimental bifurcations including the coexistence of the chest and falsetto registers as well as chaotic episodes near the register transitions. These simulations indicate that highly complex laryngeal vibrations can be traced back to instabilities of low-dimensional nonlinear dynamics.

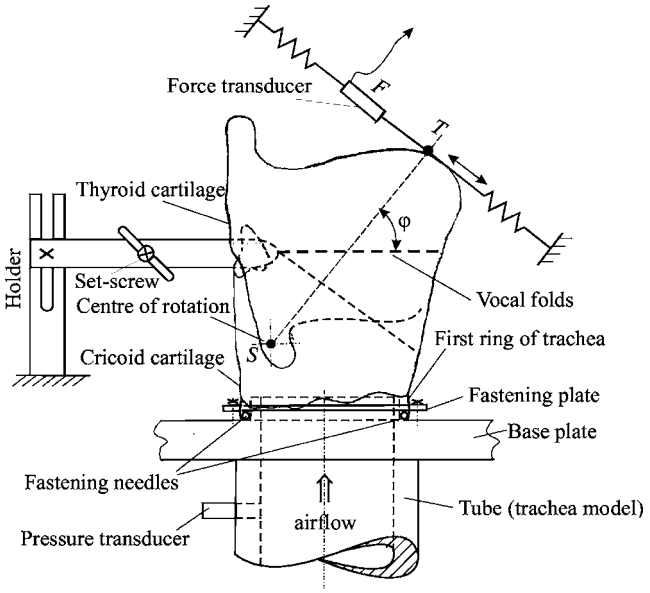

FIG. 1. Schematic illustration of the experimental setup. Human male larynge is fixed horizontally to a plate, through which the airflow is delivered from a tube to the vocal folds. Longitudinal tension of the vocal folds is monitored by a force transducer fixed to thyroid cartilage by a string. The tension is smoothly increased or decreased by rotating the thyroid cartilage with respect to cricoid cartilage.

The present paper is organized as follows. Section II introduces details of the experimental data recorded from the excised larynx experiment. The experimental data are then modeled by the nonlinear prediction technique in Sec. III. Section IV compares nonlinear dynamical properties of the biomechanical model with those of the experiment. The last section is devoted to conclusions and discussion.

II. EXPERIMENTAL DATA

Following the basic experimental procedure of van den Berg,[13,14,22] the experimental setup was prepared as illustrated in Fig. 1 (see details in Refs. 18–20). Human male larynges were fixed horizontally to a plate, through which the airflow, heated to 37 °C and humidified, was delivered to the vocal folds. The air was passed through a tube, the dimensions of which correspond to the volume of the human subglottal space. Since no vocal tract was included, the air was expelled into a free atmosphere. The airflow rate was kept constant (0.4 l/s). Subglottal pressure in the tube was measured by pressure transducers. Longitudinal tension of the vocal folds, monitored by a force transducer, was smoothly increased or decreased by maintaining a symmetric configuration of the right and left vocal folds. Changes of the fundamental phonation frequency and of the vibration regimes were measured with a microphone. The vocal fold oscillations were also observed optically by stroboscopy and videokymography.[23]

Of particular interest in this experiment is the controlled variation of the vocal fold tension, which induces abrupt

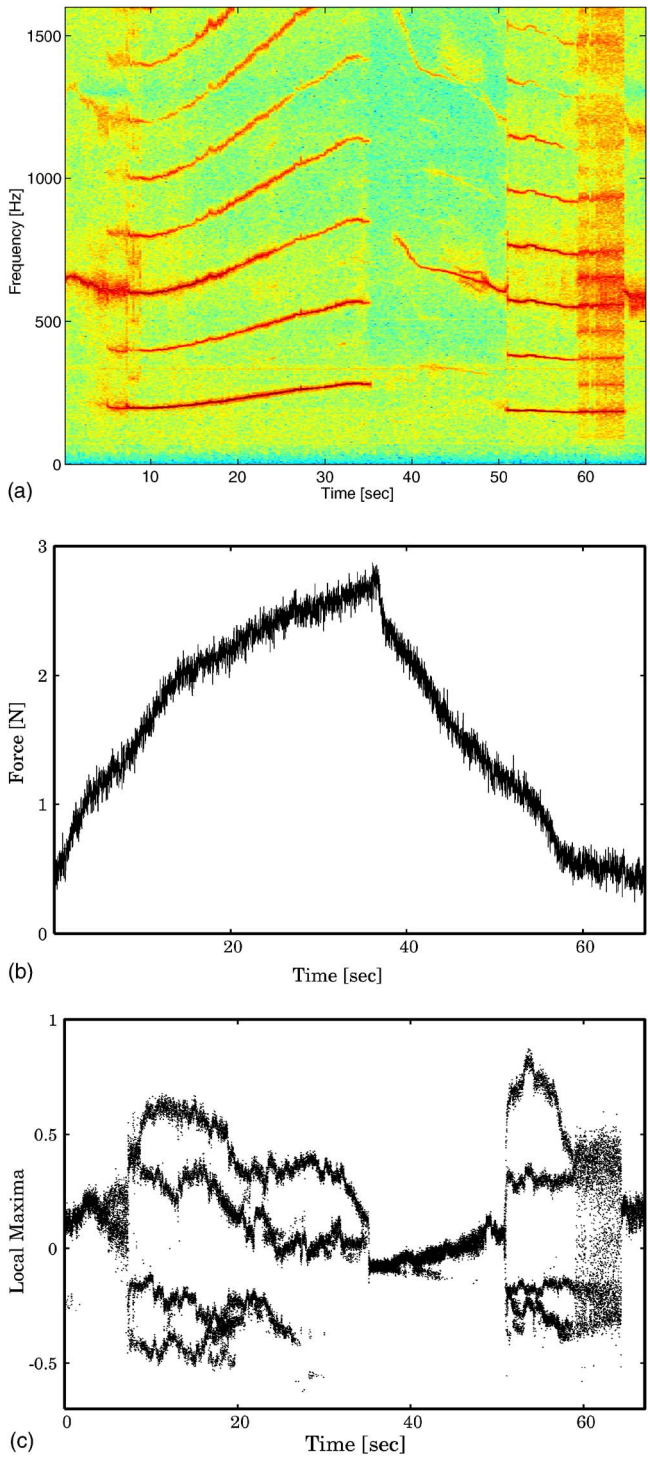

FIG. 2. (Color online) (a) Spectrogram of the microphone signal recorded from the excised human larynx. (b) Force signal that controls the vocal fold elongation. (c) Plot of a sequence of local maxima $\{x_l(n)\}$ extracted from the microphone signal of (a).

jumps between different vibration patterns. Such sudden jumps are associated with transitions from nearly harmonic vibrations without glottal closure to more intense oscillations with complete glottal closure. These vibration patterns closely resemble the falsetto and chest registers in the human voice.[24,25] Thus, careful excised larynx experiments can elucidate the dynamical basis of the voice registers.

Figures 2(a) and 2(b) display a spectrogram of the microphone signal $\{x(t)\}$, which is normalized between $[-1,1]$ (sampling frequency: 8192 Hz) and the slowly varying elongation force $\{y(t)\}$. The tension is monotonically and slowly increased for $t=37$ s and then monotonically decreased back to the smallest value at $t=67$ s. The first 5 s of the spectrogram represent falsetto-like vibrations with a fundamental frequency of about 600 Hz. At around $t=5$ s, a sudden transition to chest-like vibration occurs. This chest regime lasts until it is interrupted by a short aphonic episode at $t=35$. Backward transition to falsetto is observed around $t=38$ s. This falsetto jumps back again to chest-like vibrations around $t=51$ s. At about $t=59$ s, irregular vibrations appear until it switches to falsetto-like vibration. The described falsetto-chest transitions are found at different values of the increased or decreased elongation force, implying that the transitions exhibit hysteresis caused by the coexistence of the falsetto and chest vibrations.[14,20]

Figure 2(c) shows plots of a sequence of local maxima of the microphone signal. To draw this graph, noisy components of the microphone signal are reduced by applying a moving average filter as $x'(t)=(1/4)\Sigma_{i=0}^{3}x(t+i)$. The local maxima $x'(t)$ that satisfy the conditions of $x'(t-1)<x'(t)$ and $x'(t)<x'(t+1)$ are then extracted and are denoted in the order n of their appearance as $\{x_l(n):n=1,2,\ldots\}$. Considering the local maximum as the Poincaré section, this graph corresponds to the bifurcation diagram of periodic points. Although the branches are rather bold due to noise inherent in the experiment, clear bifurcations are recognized. The falsetto and chest regimes give rise to limit cycles with one or multiple maxima, respectively, whereas the irregular regime shows no clear periodicity.

Now we study the vibratory patterns in more detail. Figure 3 shows the microphone signal in two-dimensional delay coordinate representation $\{x(t),x(t-\tau)\}$ $(\tau=3/8192$ s).[26,27] The graphs (a), (b), and (c) correspond to falsetto, chest, and noisy oscillations observed at $t\in[49,50]$, $[54,55]$, $[62.5,63.5]$, respectively. Both falsetto and chest regimes show clear limit cycle structures. Compared with the falsetto oscillation which has an almost sinusoidal structure, the chest oscillation shows more complex dynamics related to its richer harmonic structure. The noisy oscillations, on the other hand, show more complex dynamics, which can be confirmed in the return plots of the local maxima $\{x_l(n-1),x_l(n)\}$. Whereas the return plots are densely concentrated around few areas in the falsetto and chest, the return points are widely scattered for the noisy dynamics [see Figs. 3(d)–3(f)].

III. NONLINEAR MODELING

As we have seen with the delay-coordinate representation, the chest and falsetto vibrations observed from the excised larynx experiment give rise to relatively simple limit cycle attractors. The chest-falsetto transition can then be interpreted as bifurcations of the limit cycles accompanied by irregular dynamics. Although the voice production involves complex interactions among aerodynamic, biomechanical, and acoustic factors, our observation shows that relatively

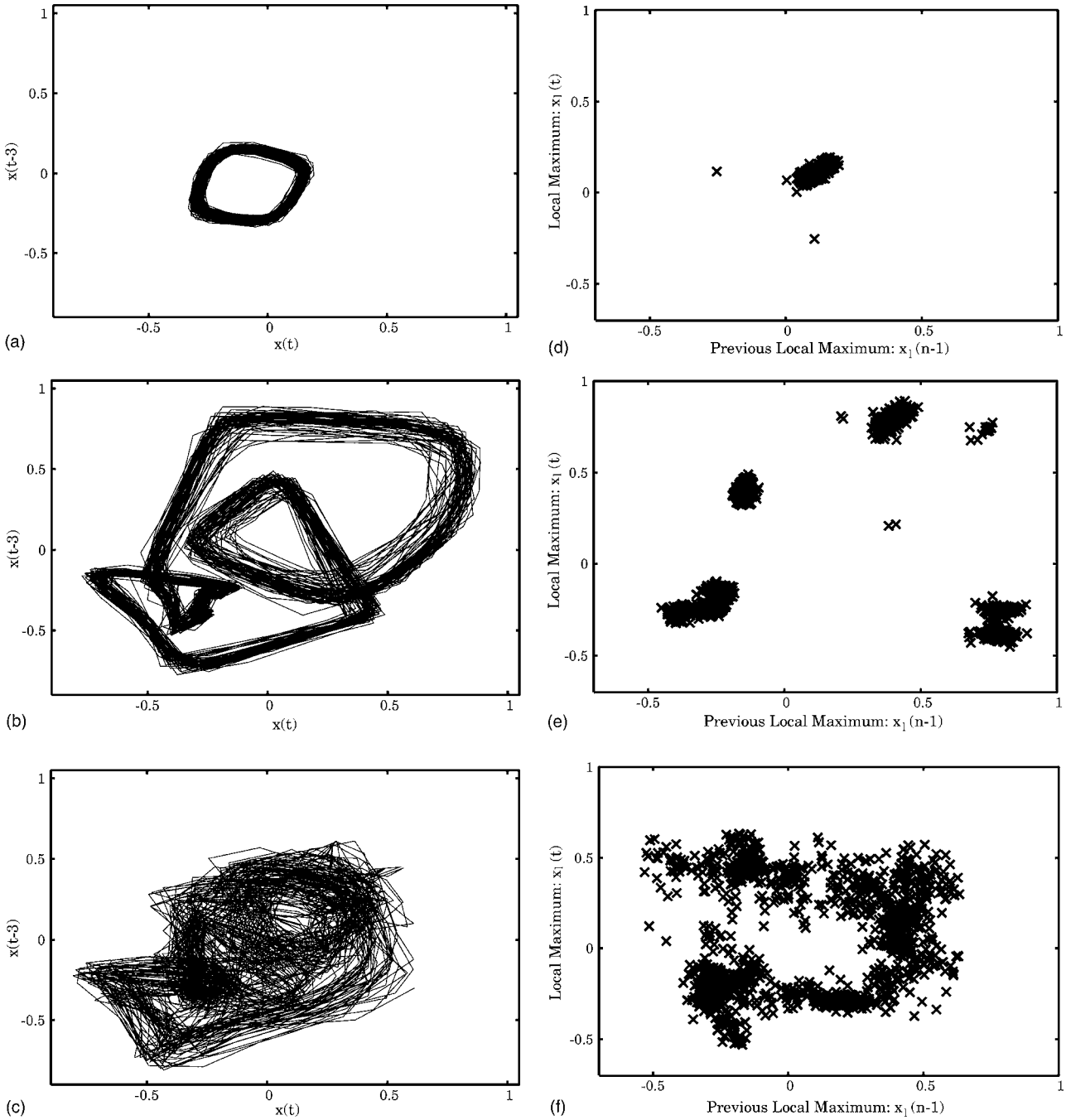

FIG. 3. (a)–(c) Delay-coordinate representation $\{x(t),x(t-\tau)\}$ ($\tau=3/8192$ s) of the microphone signal in the intervals $t\in[49,50]$, $[54,55]$, $[62.5,63.5]$, respectively. (d)–(f) Return plots $\{x_l(n-1),x_l(n)\}$ of the local maxima extracted from the corresponding intervals.

low-dimensional nonlinear dynamics underlies the excised larynx experiment. The aim of this section is to examine this low-dimensionality based upon the deterministic nonlinear modeling of the time series data. Our attempt is to show that there exists a simple deterministic nonlinear system, that can model the experimental data with high accuracy. Our modeling approach is explained as follows. First, we embed the time series $\{x(t)\}$ into delay coordinates $X(t)=\{x(t),x(t-\tau),\ldots,x(t-(d-1)\tau)\}$ (d: embedding dimension, τ: time lag). According to the embedding theorem,[26,27] there exists an associated dynamics

$$\frac{dX}{dt}=F(X(t)). \tag{1}$$

Using the Euler's formula, the above equations are discretized as $X(t+\Delta t)=X(t)+\Delta tF(X(t))$ (Δt: sampling interval). The main point of the modeling is to construct a nonlinear function $\tilde{F}$ that approximates the original F. If the experimental system is well embedded into the delay coordinate space and the model $\tilde{F}$ gives a good approximation for F, dynamics of the excised larynx experiment should be well

simulated by free-running the model $X(t+\Delta t)=X(t)+\Delta t\tilde{F}(X(t))$.

The nonlinear modeling techniques can be categorized mainly into local modeling and global modeling. The local modeling is to divide the state space into small regions and to construct a linear or nonlinear function in each region.[28] The global modeling, on the other hand, yields a single nonlinear function that approximates the global dynamics without dividing the space.[29,30] Although the local approach is capable of precisely modeling the local dynamics, it is not suitable for modeling the bifurcations that require continuous change of the global dynamical structure of the nonlinear model.[30–33] Since our aim is to retrieve bifurcations that underly the chest-falsetto transitions, the global modeling is essential for our study. As one of the most popular global techniques, we exploit the radial basis functions (RBFs).[30,33] In the RBF technique, each component of the nonlinear dynamics $\tilde{F}=\{\tilde{F}_1,\tilde{F}_2,\ldots,\tilde{F}_d\}$ is modeled as

$$\tilde{F}_i(X)=\sum_{k=1}^{K}\Omega_{i,k}\phi(\sigma_k,\|X-c_k\|), \tag{2}$$

which is a linear summation of the nonlinear basis functions ϕ, whose output is localized around the centroid c_k. The symbol $\|\cdot\|$ denotes the Euclidean norm. Once the number K of the basis functions and the parameter values $\{\sigma_k\}_{k=1}^{K}$ and the centroids $\{c_k\}_{k=1}^{K}$ of all the basis functions are determined, the linear coefficients $\Omega_{i,k}$ are obtained by the least-squares algorithm.

The procedure of modeling the time series $\{x(t)\}$ consists of the following steps:

(P1) The embedding dimension d and the time lag τ are determined. The embedding dimension was found by the false-nearest-neighborhood algorithm,[34] which indicates the minimum embedding dimension required to reconstruct nonlinear dynamics. For the falsetto, chest, and chaotic vibrations (shown in Fig. 3), we obtained the minimum embedding dimension of $d=3$, 5, and 5, respectively (detailed results not shown). Among the three values, the maximum of $d=5$ was chosen as the embedding dimension commonly used for all three vibrations. The time lag was set to be $\tau=3/8192$ s, which corresponds to the first zero-crossing point of the autocorrelation function[35] computed from the falsetto regime.

(P2) The centroids were set as

$$\{c_k=\hat{X}_k+\xi: k=1,2,\ldots,K\}, \tag{3}$$

using the points $\hat{X}$ randomly selected from the data as $\{\hat{X}_k\}\subset\{X(t)\}$. The noise part $\xi=\{\xi_1,\xi_2,\ldots,\xi_d\}$ is composed of a set of independent *Gaussian* noises $\xi\in N(0,\gamma^2)$, whose standard deviation γ is set to be 30% of that of the data $\{x(t)\}$. The centroids chosen in this way are called *chaperons*[33] and have been successfully applied to the nonlinear modeling of experimental string data. The number K of the basis functions should be large enough for a precise modeling, but it should not be too large to avoid the overfitting. In the present modeling, this number was empirically set as $K=500$, which was sufficient for reproducing the experimental data. Similar results can be obtained by varying K from $K=300$ to 1000.

(P3) As a basis function, the Gaussian RBF, i.e., $\phi(\sigma_k,r)=\exp(-r^2/\sigma_k^2)$, was used. The variance parameter was set nonuniformly to each basis function as $\sigma_k^2=\min_{i\neq k}\|\mathbf{c}_i-\mathbf{c}_k\|^2$, so that each basis function covers enough range of the localized state space around the centroid and that the basis functions have not much overlap with each other. The present nonuniform setting of the variance parameters has not been widely used for the RBF technique. However, this setting seems to be quite effective for the modeling of both chest and falsetto registers in the excised larynx experiment.

(P4) The linear coefficients $\{\Omega_{i,k}\}$ can be obtained by the least-squares error algorithm; that is, to minimize the cost function

$$E(\Omega)=\sum_t\left\|\frac{X(t+\Delta t)-X(t)}{\Delta t}-\tilde{F}(X(t))\right\|^2. \tag{4}$$

(P5) From a given initial condition $X(s)$ chosen from one of the experimental data, the dynamics of the excised larynx can be simulated by free-running the model equations $X(t+\Delta t)=X(t)+\Delta t\tilde{F}(X(t))$.

Figure 4 shows trajectories and return plots generated from the nonlinear model, where the graphs (a), (b), and (c) correspond to the modeling of the falsetto, chest, and irregular vibrations, respectively. It can been seen that the dynamical structure of the original vibratory pattern shown in Fig. 3 is well reproduced by the model. The falsetto is modeled as a limit cycle with a simple harmonic structure indicating a single point in the return plot (d), whereas the chest is also modeled as a limit cycle but with a much more complicated structure indicating five points in the return plot (e). The irregular vibration is described by much more complicated dynamics indicating scattered points in the return plot (f). According to the Lyapunov spectrum analysis[36] applied to the nonlinear model, the maximum Lyapunov exponent was estimated to be positive ($\lambda_1=398.0\ \mathrm{s}^{-1}$), implying that the irregular dynamics is chaotic.

Now we address the challenging case of modeling the bifurcations of the excised larynx experiment including the chest-falsetto transitions. The entire time series data $\{x(t)\}$ of the excised larynx experiment are divided into small parts $\Xi_n=\{x(t): t\in[L(n-1),Ln]\}$ (L: division length), within which the data are considered to be stationary. For each part Ξ_n, the nonlinear model $\tilde{F}_n$ was constructed according to (P1)–(P5). The bifurcation diagram can then be drawn by successively simulating the nonlinear models $\tilde{F}_n$, where the dynamics between the two successive models $(\tilde{F}_n,\tilde{F}_{n+1})$ is realized by the linear interpolation $\tilde{F}=\alpha\tilde{F}_n+(1-\alpha)\tilde{F}_{n+1}$ ($0\leq\alpha\leq1$).

Figure 5(b) shows the bifurcation diagram reconstructed by the nonlinear model. The original time series with a total length of 67.1 s was divided into 220 segments with a length $L=2500/8192$ s. Accordingly, the corresponding nonlinear models $\{\tilde{F}_n: n=1,2,\ldots,220\}$ were constructed. By simulating the nonlinear model and extracting the local maxima of

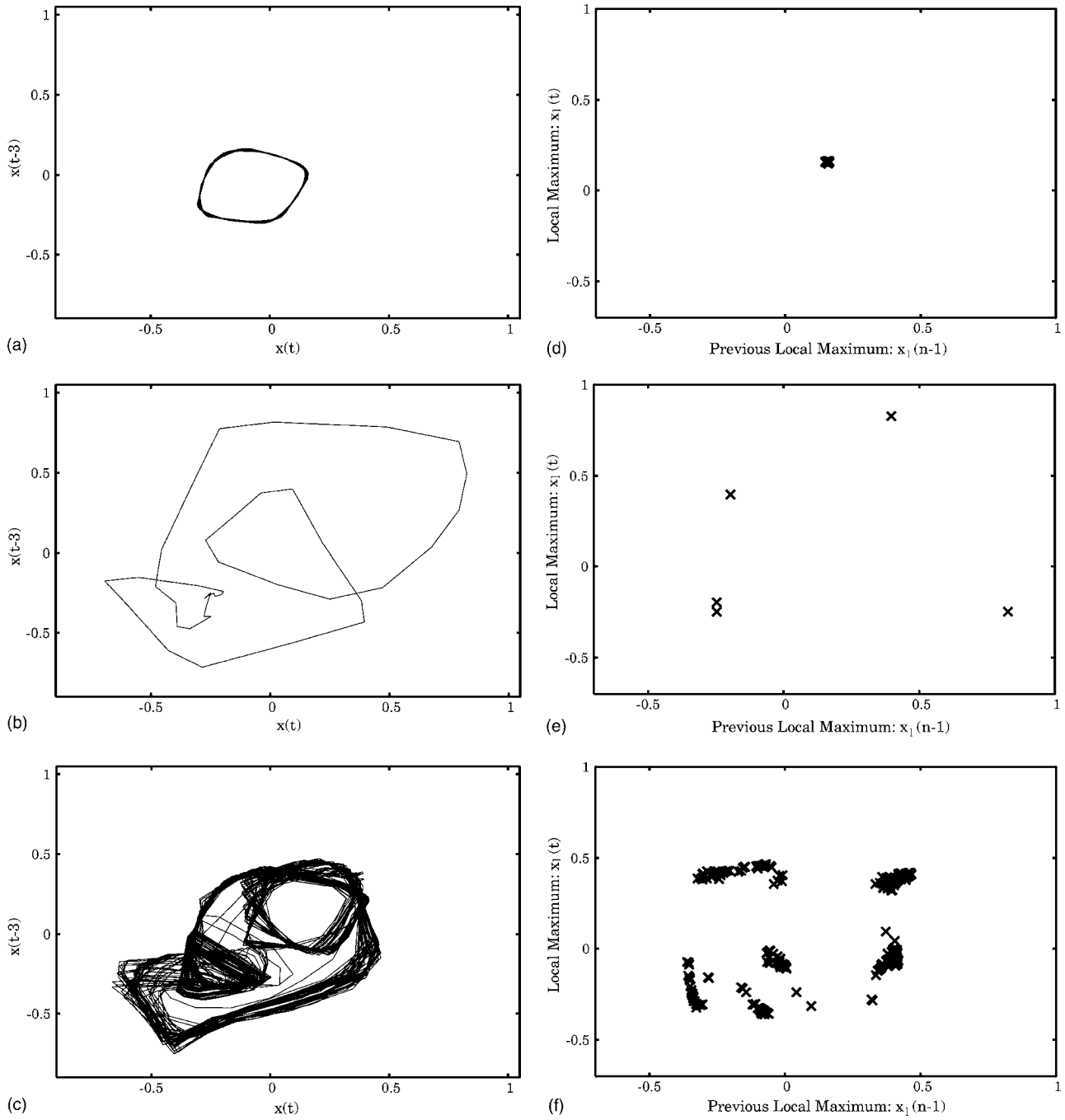

FIG. 4. (a)–(c) Dynamical trajectory in delay coordinate space generated from the nonlinear model constructed from the microphone signal between interval $t \in [49,50]$, $[54,55]$, $[62.5,63.5]$, respectively. (d)–(f) Return plots of the local maxima obtained from the same nonlinear model as (a), (b), (c), respectively.

the first component of the trajectory $X(t)$, the bifurcation diagram was drawn, showing a strong similarity to the original diagram of Fig. 2. The corresponding spectrogram of Fig. 5(a) indicates that a single branch at $t \in [0,5.4]$ and $t \in [35.5,51.3]$ corresponds to falsetto vibrations with high frequency. Multiple branches seen at $t \in [5.4,35.5]$ and $t \in [51.3,59]$ correspond to chest vibrations with lower fundamental frequency. Transitions between the chest and falsetto are accompanied by scattered points around $t \approx 6$, $t \approx 51$, and $t \in [59,64.5]$, which show broadband spectrum structure. The maximum Lyapunov exponent computed from the nonlinear model by the standard technique[36] indicates that the dynamics in these regimes are chaotic in the sense of positive Lyapunov exponent [see Fig. 5(c)]. The positive Lyapunov exponent is indicated also in other small regions such as $t \approx 18$ or 41.

In summary, the essential features of the nonlinear dynamics observed from the excised larynx experiment have been reproduced by relatively simple nonlinear models. Interpretation of the chest and falsetto vibrations as limit cycles has been justified by generating the attractors from the nonlinear models. The chest-falsetto transition accompanied by chaos was confirmed by studying the bifurcations of the nonlinear models, that produce chaotic episodes close to the

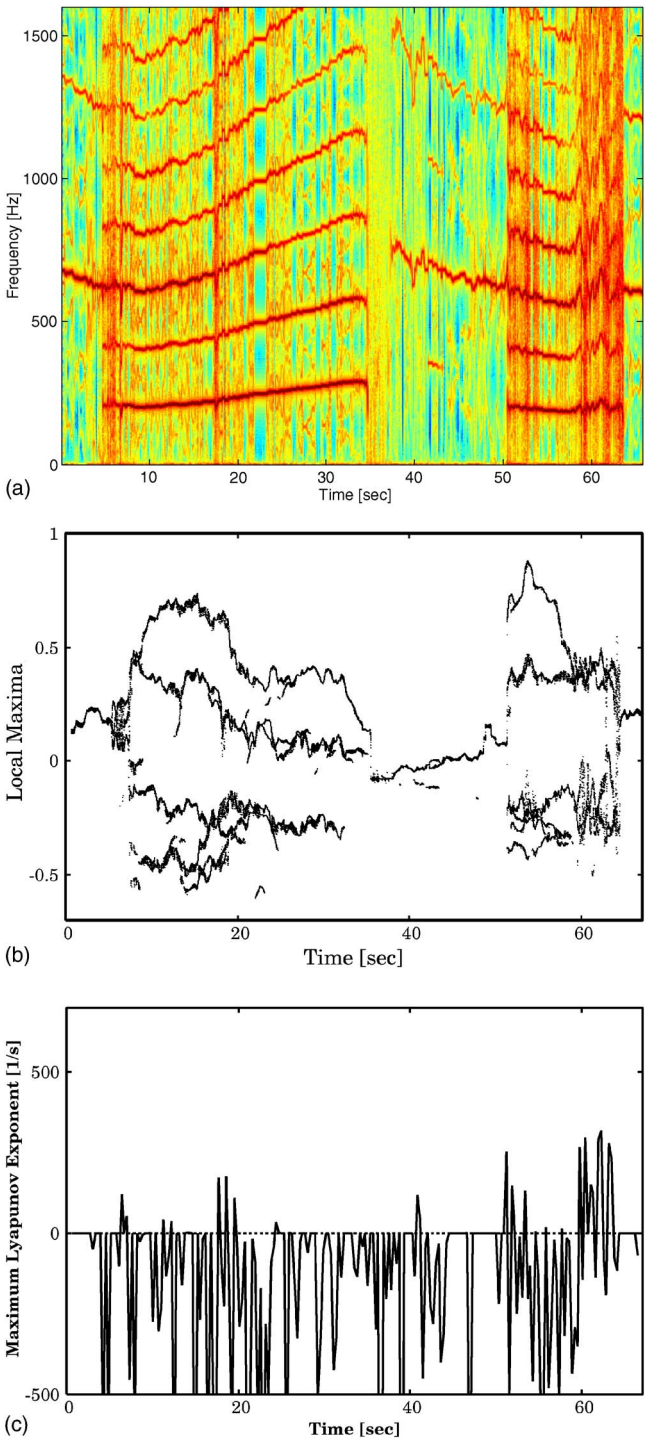

FIG. 5. (Color online) (a) Spectrogram of the dynamical signal generated from the nonlinear model. (b) Bifurcation diagram drawing a sequence of local maxima of the signal generated from the nonlinear model. (c) Maximum Lyapunov exponent computed from the nonlinear model.

transition point. Reproduction of the qualitatively very similar dynamics implies that the excised larynx experiment is well elucidated in terms of deterministic low-dimensional nonlinear dynamics.

We have looked at the register transition in several experimental data using other human excised larynges. Although the observed nonlinear phenomena can be variable, it was one of the typical observations that the chest-falsetto transition is accompanied by chaotic regime. For instance, Fig. 6 shows an example from another subject. In this experiment, the elongation force was initially set to 2 N, monotonically decreased to 0.8 N until $t \approx 10$ s, and then monotonically increased back to 2 N [see graph (b)]. The graphs (a), (c), and (d) show a spectrogram of the microphone signal from the experiment, bifurcation diagram that plots local maxima of the same microphone signal, and bifurcation diagram reconstructed by the nonlinear model. The modeling procedure described above was applied with $d=5$, $\tau = 3/8192$ s, $K=400$, and $L=1000/8192$ s. Again, the single branch with frequency 600 Hz corresponds to falsetto, whereas the multiple branches with 150 Hz correspond to chest. The chest-falsetto transition is accompanied by a chaotic episode at $t \in [13.5, 17.5]$. The bifurcation structure has been very well reproduced by the nonlinear model. This implies that our observation seems rather common in excised larynx experiments.

IV. BIOMECHANICAL MODELING

In the previous section, nonlinear modeling was applied to excised larynx data. This has been a top-down approach, which does not take into account the physics underlying the experiment. The aim of this section is to provide further insight into the experiment by introducing a biomechanical model that considers basic oscillatory mechanisms of the vocal folds. In the study of voice production, numerous models have been developed for the vocal fold oscillations ranging from simplified low-dimensional models[37–40] to complex high-dimensional models.[41] There exist, however, only a few models that simulate the transitions between chest and falsetto registers. To simulate such register transitions, we have recently introduced a three-mass model that elucidates the excised larynx experiments.[42] The significant feature of this model is the coexistence of the chest and falsetto registers within the same dynamical configuration. Comparison of its spectral characteristics with those of the experiments justified the model. Here we compare nonlinear dynamical characteristics such as Lyapunov exponent, phase trajectories, and return plots of the three-mass model with those of the experiment.

The three-mass model[42] was constructed by adding one more mass on top of the two-mass model[37] (see Fig. 7). Addition of the third mass divides the upper part of the vocal folds into two portions, which can vibrate out of phase. These phase differences can simulate the mucosal waves, which are observed in the videokymograms of both chest and falsetto registers.[15,19,43] In particular, during the falsetto register, the waves are visible only on the thin upper medial portion of the vocal folds and on the upper vocal fold surface. In the present three-mass model, such an oscillatory mode can be realized by the anti-phase oscillations between the upper and the middle masses.[42] Compared with the two-mass model, which was not designed to model such small oscillations of the upper vocal fold, the three-mass model has the advantage of modeling the falsetto register based on the upper vocal fold oscillations, which can easily coexist with the chest register. Our main modeling assumptions are (1)

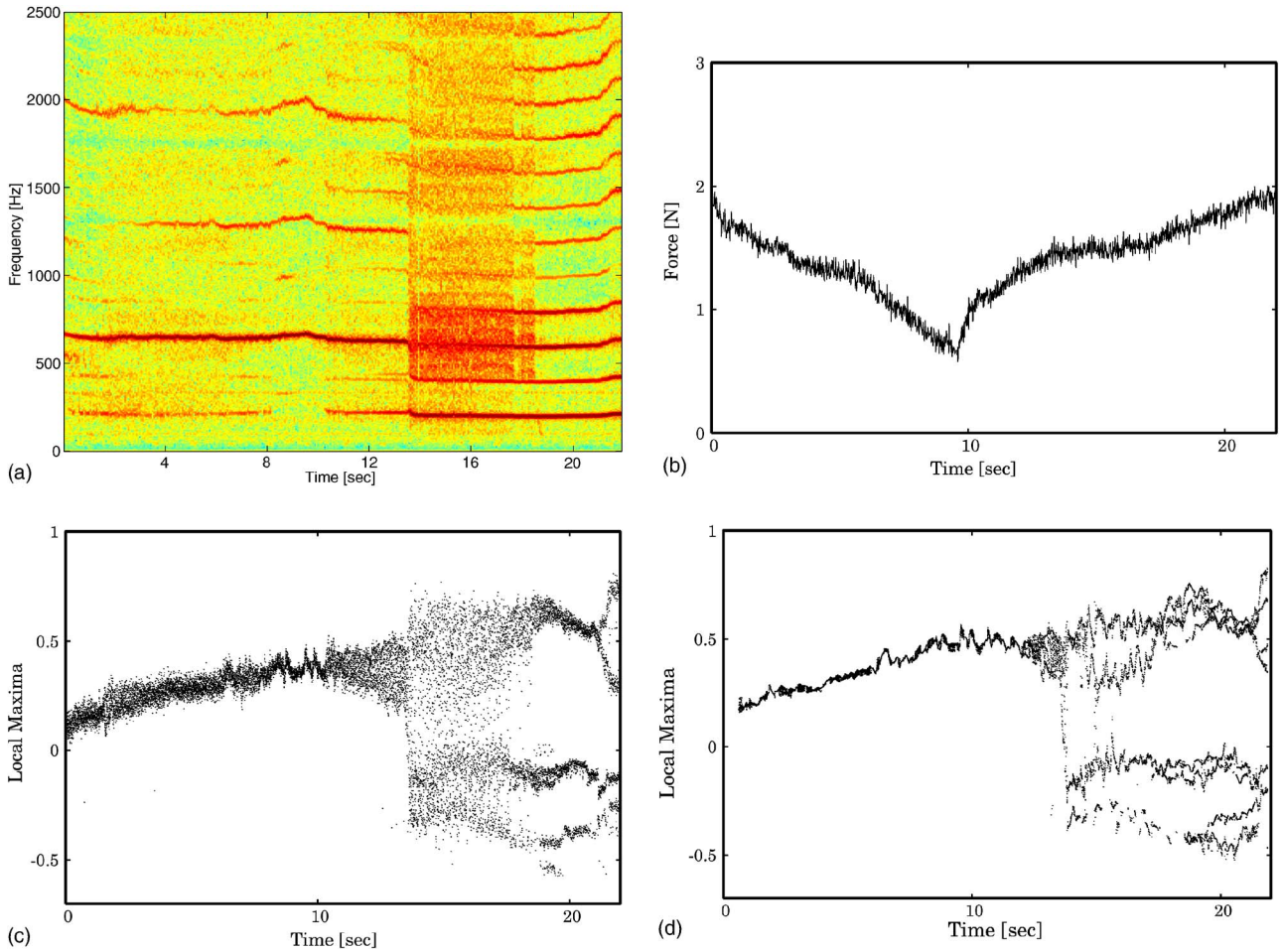

FIG. 6. (Color online) (a) Spectrogram of the microphone signal recorded from the excised larynx experiment of another subject. (b) Force signal that controls the vocal fold elongation. (c) Plot of a sequence of local maxima extracted from the microphone signal of (a). (d) Plot of a sequence of local maxima extracted from the dynamical signal generated from the nonlinear model.

the three masses are coupled by linear springs, (2) the air flow inside the glottis is described by the Bernoulli's principle below the narrowest part of the glottis,[6] (3) there is no influence of the vocal tract and the subglottal resonances (as in the experiment), and (4) the left and the right vocal folds are symmetric to each other. The model equations then read

$$m_1\ddot{x}_1 + r_1\dot{x}_1 + k_1x_1 + \Theta(-a_1)c_1\left(\frac{a_1}{2l}\right) + k_{1,2}(x_1 - x_2) = ld_1P_1, \tag{5}$$

$$m_2\ddot{x}_2 + r_2\dot{x}_2 + k_2x_2 + \Theta(-a_2)c_2\left(\frac{a_2}{2l}\right) + k_{1,2}(x_2 - x_1) + k_{2,3}(x_2 - x_3) = ld_2P_2, \tag{6}$$

$$m_3\ddot{x}_3 + r_3\dot{x}_3 + k_3x_3 + \Theta(-a_3)c_3\left(\frac{a_3}{2l}\right) + k_{2,3}(x_3 - x_2) = 0. \tag{7}$$

The dynamical variables x_i represent displacements of the masses m_i (lower mass: $i=1$, middle mass: $i=2$, upper mass: $i=3$), where the corresponding glottal areas are given by $a_i = a_{0i} + 2lx_i$ (a_{0i}: prephonatory area, l: length of the glottis). The constant parameters r_i, k_i, c_i, and d_i represent damping, stiffness, collision, and thickness of the masses m_i,

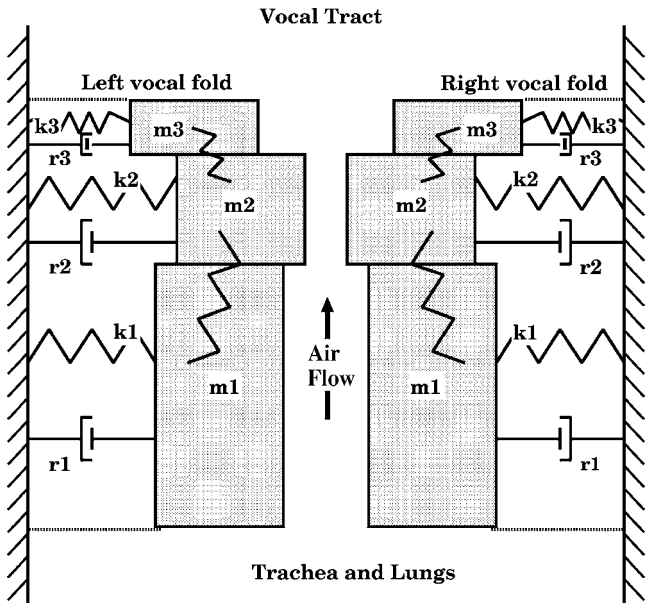

FIG. 7. Schematic illustration of the three-mass model. The left and right vocal folds have a symmetric configuration. Each vocal fold is composed of three masses (m_1: lower mass, m_2: middle mass, m_3: upper mass) coupled by linear springs. The air flow coming from the lungs is described by the Bernoulli's principle below the narrowest part of the glottis.

respectively, whereas $k_{i,j}$ represents coupling strength between two masses m_i and m_j. The collision function is approximated as $\Theta(x)=0$ $(x\leq 0)$; $\Theta(x)=\tanh(1000x)$ $(0<x)$.[6] The pressures P_i, which act on the masses m_i, are given by $P_1=P_s[1-\Theta(a_{\min})(a_{\min}/a_1)^2]\Theta(a_1)$, $P_2=P_s[1-\Theta(a_{\min})(a_{\min}/a_2)^2]\Theta(a_1)\Theta(a_2)\Theta(a_1-a_3)\Theta(a_2-a_3)$, where $a_{\min}=\min(a_1,a_2,a_3)$. The glottal volume flow velocity is computed as $U=\sqrt{(2P_s/\varrho)}a_{\min}\Theta(a_{\min})$ (ϱ: air density).[6] The tension parameter $\tilde{Q}$ is defined to determine the size and the stiffness of the second mass as follows:

$$k_2 = 0.08\tilde{Q}\ [\mathrm{g/ms^2}], \tag{8}$$

$$m_2 = 0.025/\tilde{Q}\ [\mathrm{g}]. \tag{9}$$

In the sense that the natural frequency of the second mass is roughly given by $f_2=(1000/2\pi)\sqrt{k_2/m_2}=284.7\tilde{Q}$ Hz, the tension parameter linearly controls the frequency of the second mass. To simulate the excised larynx experiment, this tension parameter $\tilde{Q}$ is changed as the main bifurcation parameter. The other parameter values, which are summarized in the Appendix, were adopted from the standard values established in the two-mass models.[6,37]

Figure 8(b) displays bifurcation diagram of the three-mass model. The bifurcation parameter was changed bidirectionally in $\tilde{Q}\in[0.9,1.58]$. When reversing the changing direction at $\tilde{Q}=1.58$, the initial condition was set to be a stable fixed point to emphasize the existence of an aphonic episode. To see the corresponding frequency structure, the spectrogram was drawn in Fig. 8(a) by increasing the bifurcation parameter $\tilde{Q}\in[0.9,1.58]$ from $t=0$ s to $t=50$ s, and then by decreasing it from $t=50$ s to $t=100$ s. This resembles the spectrogram of the real experimental data of Fig. 2(a) quite well. In the increasing direction of $\tilde{Q}$, low-frequency oscillations dominate the bifurcation diagram, whereas in the decreasing direction of $\tilde{Q}$, higher-frequency oscillations last until they switch to low-frequency oscillations at $\tilde{Q}\approx 1.03$ (indicated by an arrow). This hysteresis is clearly due to the coexistence of the low-frequency and high-frequency oscillations, which correspond, respectively, to chest and falsetto vibrations according to the vibratory patterns observed by the simulation. Complete glottal closure was observed for the chest-like vibrations, whereas glottal area was not completely closed for the falsetto-like vibrations. Irregular dynamics are observed in two regimes: one is in the chest-like vibration regime at $\tilde{Q}\in[1.05,1.13]$ and the other is during the register transition from falsetto to chest around $\tilde{Q}\approx 1.03$. In Fig. 8(c), the maximum Lyapunov exponent was computed by the standard algorithm.[36] The two regimes are indeed chaotic with a positive Lyapunov exponent, which indicates its orbital instability.

Figure 9 shows three attractors in $(U,\dot{U})$-space generated from the three-mass model, where the graphs (a), (b), and (c) correspond to $\tilde{Q}=1$, 1.1, 1.2, respectively. The parameter $\tilde{Q}=1$ gives rise to a chest-like limit cycle (139.5 Hz) with five points in the return plot (d), whereas $\tilde{Q}=1.2$ produces a falsetto-like limit cycle (315 Hz) with only two points closely located to each other in the return plot (f). Between the chest-like and falsetto-like limit cycles, the parameter $\tilde{Q}=1.1$ generates chaotic oscillations with many

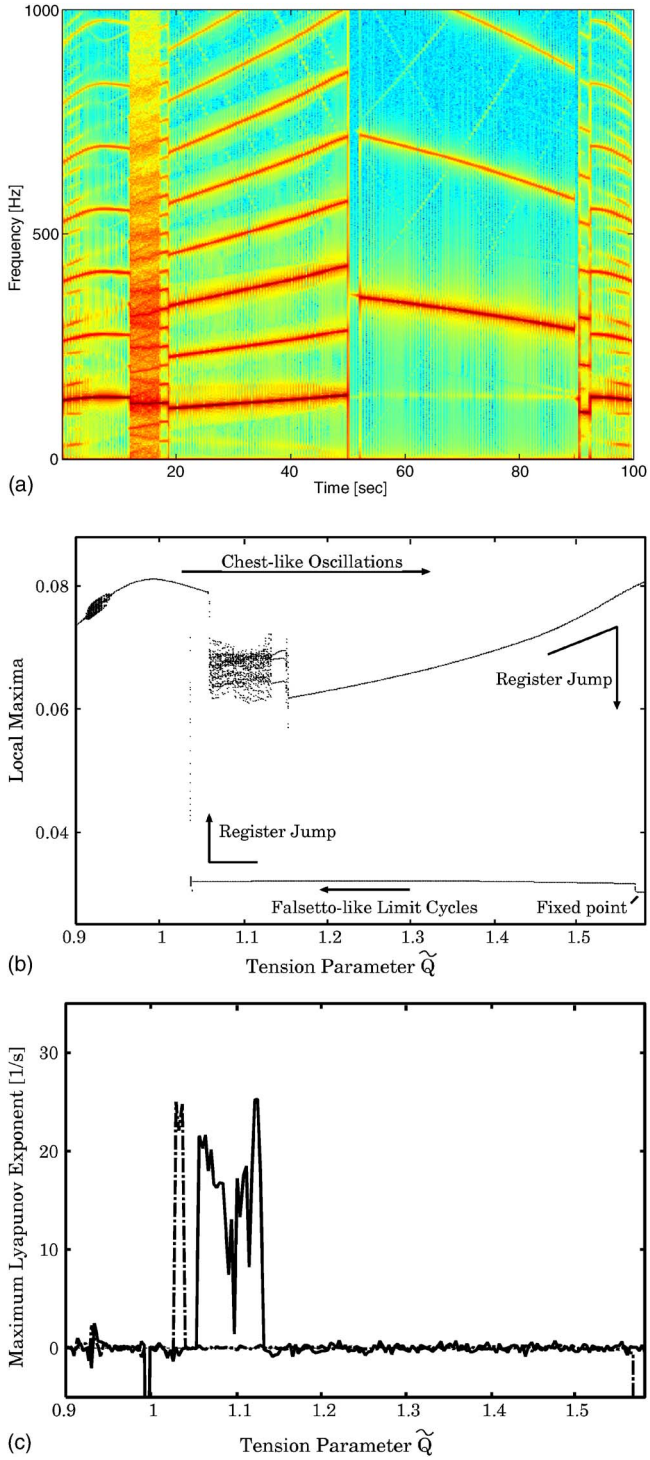

FIG. 8. (Color online) (a) Spectrogram of the dynamical signal $\{x_1(t)\}$ generated from the three-mass model. The tension parameter is increased from $\tilde{Q}=0.9$ to $\tilde{Q}=1.58$ in $t\in[0,50]$ and then decreased back to $\tilde{Q}=0.9$ in $t\in[50,100]$. (b) Bifurcation diagram drawing the local maxima of variable $x_1(t)$ from the three-mass model. The tension parameter is changed bidirectionally in $\tilde{Q}\in[0.9,1.58]$. (c) Maximum Lyapunov exponent computed for the three-mass model. The solid and dot-dashed lines correspond to the increasing and decreasing directions of $\tilde{Q}$, respectively.

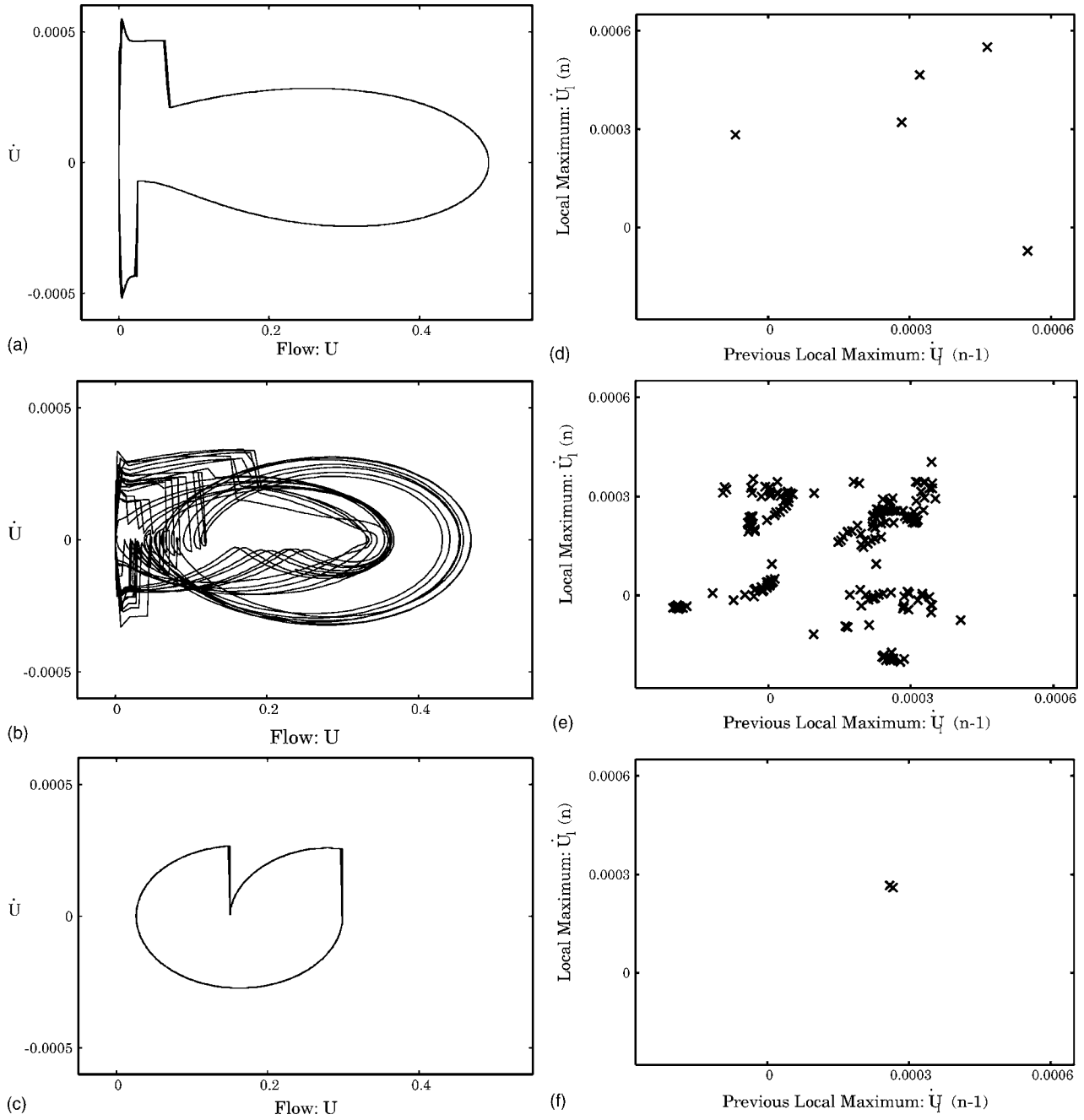

FIG. 9. (a)–(c) Dynamical trajectory in $(U,\dot{U})$-space generated from the three-mass model with $\tilde{Q}=1,\ 1.1,\ 1.2$, respectively. (d)–(f) Return plots of the local maxima obtained from variable $\dot{U}(t)$ of the three-mass model corresponding to (a), (b), (c), respectively.

scattered points in the return plot (e). Comparing with the attractors reconstructed from the experimental data in Fig. 3, the presented three cases reproduce the qualitative structure of the excised larynx experiment.

The presented bifurcation analysis demonstrates that the three-mass model shows good agreement with the excised larynx experiment. It reproduces the coexistence of the chest and falsetto oscillations and abrupt transitions between them. During such transitions, the dynamics goes through chaotic regimes between the chest and falsetto oscillations.

V. CONCLUSIONS AND DISCUSSION

Data from the excised larynx experiment have been analyzed by nonlinear data analysis. A main feature of this experiment is the coexistence of the two distinct vibration patterns that resemble the chest and falsetto registers of human voice. Nonlinear modeling technique clarifies that these chest-like and falsetto-like vibrations can be characterized as limit cycles. Abrupt transitions between the two vibration patterns correspond to jumps between the coexisting limit cycles typically accompanied by chaotic dynamics, which can be reconstructed in five-dimensional delay-coordinate space. Biomechanical modeling shows further evidence that the low-dimensional modeling is indeed valid for the excised larynx experiment.

In the modeling studies of the vocal fold oscillations, the main mechanisms known to generate deterministic chaos are left-right asymmetry,[6] anterior-posterior modes,[44] exces-

sively high subglottal pressure,[21] acoustical coupling with supraglottal resonances,[45] or additional vibrating tissues such as vocal membranes.[11] In particular, left-right asymmetry and excessively high subglottal pressure have been confirmed to generate chaos in excised larynx experiments.[15,17] This study presents a new mechanism of how chaos can be induced near the chest-falsetto transition of vocal fold vibrations. Of particular interest in future studies is the analysis of bifurcations for register transitions of human subjects quantified by high-speed recordings[46,47] or videokymography.[23]

ACKNOWLEDGMENTS

This work was supported by the Alexander-von-Humboldt Foundation, the Deutsche Forschungsgemeinschaft, SCOPE (071705001) of Ministry of Internal Affairs and Communications (MIC) of Japan, and the Grant Agency of the Academy of Sciences of the Czech Republic by Project No. IAA2076401 "Mathematical modeling of human vocal fold oscillations."

APPENDIX: MODEL PARAMETERS AND SIMULATION DETAILS

Here, the parameter values used to simulate the three-mass model in Sec. IV are listed: $m_1=0.125$ g; $m_3=0.005$ g; $d_1=0.25$ cm; $d_2=0.05$ cm; $d_3=0.01$ cm; $r_1=0.01$ g/ms; $r_2=0.005$ g/ms; $r_3=0.005$ g/ms; $k_1=0.08$ g/ms^2; $k_3=0.01$ g/ms^2; $k_{1,2}=0.05$ g/ms^2; $k_{2,3}=0.01$ g/ms^2; $c_1=3k_1$; $c_2=3k_2$; $c_3=3k_3$; $a_{01}=0.01$ cm^2; $a_{02}=0.01$ cm^2; $a_{03}=0.01$ cm^2. $l=1.4$ cm; $\varrho=0.00113$ g/cm^3; $P_s=0.008$ g/cm ms^2.

These parameter values have been adopted mainly from the standard parameters established in the two-mass models.[6,37] To observe a pronounced regime of coexistence of chest and falsetto registers, prephonatory areas and some damping constant are reduced from the standard values. The mass size and the stiffness of the second mass are determined by the tension parameter $\tilde{Q}$ according to Eqs. (8) and (9).

The initial values for all simulations were set as $x_1=x_2=x_3=0.1$ and $\dot{x}_1=\dot{x}_2=\dot{x}_3=0$. To integrate the three-mass model equations (5)–(7), the fourth-order Runge–Kutta method was applied with an integration step of $\Delta t=1/1000$ ms. A smaller integration step was tried to confirm that essentially the same results can be obtained.

The spectrograms were computed using the x_1 component of the three-mass model with the following parameters. Sampling rate: 4 kHz; window length: 2048 sample points; overlap: 0; windowing: Hanning. Note that other components such as $x_{2,3}$, $\dot{x}_{1,2,3}$, or glottal flow velocity U give similar spectrograms.

[1] I. R. Titze, *Principles of Voice Production* (Prentice-Hall, Englewood Cliffs, NJ, 1994).
[2] H. Herzel, D. Berry, I. R. Titze, and I. Steinecke, "Nonlinear dynamics of the voice," Chaos **5**, 30 (1995).
[3] I. Tokuda and H. Herzel, "Detecting synchronizations in an asymmetric vocal fold model from time series data," Chaos **15**, 013702 (2005).
[4] I. Steinecke and H. Herzel, "Bifurcations in an asymmetric vocal fold model," J. Acoust. Soc. Am. **97**, 1571 (1995).
[5] I. R. Titze, R. J. Baken, and H. Herzel, "Evidence of chaos in vocal fold vibration," in *Vocal Fold Physiology*, edited by I. R. Titze (Singular, San Diego, 1993), pp. 143–188.
[6] H. Herzel, D. Berry, I. R. Titze, and S. Saleh, "Analysis of vocal disorders with methods from nonlinear dynamics," J. Speech Hear. Res. **37**, 1008 (1994).
[7] J. Neubauer, E. Edgerton, and H. Herzel, "Nonlinear phenomena in contemporary vocal music," J. Voice **18**, 1 (2004).
[8] I. Wilden, H. Herzel, G. Peters, and G. Tembrock, "Subharmonics, biphonation, and deterministic chaos in mammal vocalization," Bioacoustics **9**, 171 (1998).
[9] M. S. Fee, B. Shraiman, B. Pesaran, and P. P. Mitra, "The role of nonlinear dynamics of the syrinx in the vocalizations of a songbird," Nature (London) **395**, 67 (1998).
[10] T. Fitch, H. Herzel, J. Neubauer, and M. Hauser, "Calls out of chaos: The adaptive significance of nonlinear phenomena in primate vocal production," Anim. Behav. **63**, 407 (2002).
[11] P. Mergell, W. T. Fitch, and H. Herzel, "Modeling the role of non-human vocal membranes on phonation," J. Acoust. Soc. Am. **105**, 2020 (1999).
[12] I. Tokuda, T. Riede, J. Neubauer, M. J. Owren, and H. Herzel, "Nonlinear analysis of irregular animal vocalizations," J. Acoust. Soc. Am. **111**, 2908 (2002).
[13] J. van den Berg, "Sound productions in isolated human larynges," Ann. N.Y. Acad. Sci. **155**, 18 (1968).
[14] J. G. Švec, H. K. Schutte, and D. G. Miller, "On pitch jumps between chest and falsetto registers in voice: Data from living and excised human larynges," J. Acoust. Soc. Am. **106**, 1523 (1999).
[15] D. A. Berry, H. Herzel, I. R. Titze, and B. H. Story, "Bifurcations in excised larynx experiments," J. Voice **10**, 129 (1996).
[16] J. J. Jiang, Y. Zhang, and C. N. Ford, "Nonlinear dynamics of phonations in excised larynx experiments," J. Acoust. Soc. Am. **114**, 2198 (2003).
[17] Y. Zhang and J. J. Jiang, "Spatiotemporal chaos in excised larynx vibrations," Phys. Rev. E **72**, 035201(R) (2005).
[18] J. Horáček, J. G. Švec, J. Veselý, and E. Vilkman, "Experimental study of the vocal fold vibration in excised larynx," *Proceedings of Interaction and Feedbacks 2000*, edited by I. Zolotarev (Institute of Thermomechanics, Academy of Sciences, Prague, 2000), pp. 27–34.
[19] J. Horáček, J. G. Švec, J. Veselý, E. Vilkman, I. Klepáček, and A. Vetešník, "Measurement of the vocal-fold vibration behaviour in excised human larynges," *Proceedings of 2nd International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications* (University of Firenze, Depts. of Electronics and Physics, Firenze, Italy, 2001), CD-ROM.
[20] J. Horáček, J. G. Švec, J. Veselý, and E. Vilkman, "Bifurcations in excised larynges caused by vocal fold elongation," *Proceedings of the International Conference on Voice Physiology and Biomechanics*, edited by A. Giovanni, P. Dejonckere, M. Ouaknine (Laboratory of Audio-Phonology, Marseille, 2004), pp. 87–89.
[21] J. J. Jiang, Y. Zhang, J. Stern, "Modeling of chaotic vibrations in symmetric vocal folds," J. Acoust. Soc. Am. **110**, 2120 (2001).
[22] E. Vilkman, "An apparatus for studying the role of the cricothyroid articulation in the voice production of excised human larynges," Folia Phoniatr. **39**, 169 (1987).
[23] J. G. Švec and H. K. Schutte, "Videokymography: High-speed line scanning of vocal fold vibration," J. Voice **10**, 201 (1996).
[24] M. Hirano, W. Vennard, and J. Ohala, "Regulation of register, pitch, and intensity of voice," Folia Phoniatr. **22**, 1 (1970).
[25] M. Hirano and Y. Kakita, "Cover-body theory of vocal cord vibration," in *Speech Science*, edited by R. G. Daniloff (College Hill Press, San Diego, 1985), pp. 1–46.
[26] F. Takens, "Detecting strange attractors in turbulence," Lect. Notes Math. **898**, 366 (1981).
[27] T. Sauer, J. A. York, and M. Casdagli, "Embeddology," J. Stat. Phys. **65**, 579 (1991).
[28] J. D. Farmer and J. J. Sidorowich, "Predicting chaotic time series," Phys. Rev. Lett. **59**, 845 (1987).
[29] J. P. Crutchfield and B. S. McNamara, "Equations of motion from a data series," Complex Syst. **1**, 417 (1987).
[30] M. Casdagli, "Nonlinear prediction of chaotic time series," Physica D **35**, 335 (1989).
[31] R. Tokunaga, S. Kajiwara, and T. Matsumoto, "Reconstructing bifurcation diagrams only from time waveforms," Physica D **79**, 348 (1994).
[32] I. Tokuda, S. Kajiwara, R. Tokunaga, and T. Matsumoto, "Recognizing

chaotic time-waveforms in terms of a parametrized family of nonlinear predictors," Physica D **95**, 380 (1996).
[33]K. Judd and A. I. Mees, "Modeling chaotic motions of a string from experimental data," Physica D **92**, 221 (1996).
[34]M. B. Kennel, R. Brown, and H. D. I. Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geometric construction," Phys. Rev. A **45**, 3403–3411 (1992).
[35]H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis* (Cambridge University Press, Cambridge, 1997).
[36]I. Shimada and T. Nagashima, "A numerical approach to ergodic problem of dissipative dynamical systems," Prog. Theor. Phys. **61**, 1605 (1979).
[37]K. Ishizaka and J. L. Flanagan, "Synthesis of voiced sounds from a two-mass model of the vocal cords," Bell Syst. Tech. J. **51**, 1233 (1972).
[38]X. Pelorson, A. Hirschberg, R. R. van Hassel, A. P. J. Wijnands, and Y. Auregan, "Theoretical and experimental study of quasi-steady flow separation within the glottis during phonation," J. Acoust. Soc. Am. **96**, 3416 (1994).
[39]B. H. Story and I. R. Titze, "Voice simulation with a body-cover model of the vocal folds," J. Acoust. Soc. Am. **97**, 1249 (1995).
[40]J. Horáček, P. Šidlof, and J. G. Švec, "Numerical simulation of self-oscillations of human vocal folds with Hertz model of impact forces," J. Fluids Struct. **20**, 853 (2005).
[41]F. Alipour, D. A. Berry, and I. R. Titze, "A finite-element model of vocal-fold vibration," J. Acoust. Soc. Am. **108**, 3003 (2000).
[42]I. T. Tokuda, J. Horáček, J. G. Švec, and H. Herzel, "Comparison of biomechanical modeling of register transitions and voice instabilities with excised larynx experiments," J. Acoust. Soc. Am. **122**, 519 (2007).
[43]J. G. Švec, F. Šram, and H. K. Schutte, "Videokymography in voice disorders: What to look for?" Ann. Otol. Rhinol. Laryngol. **116**, 172 (2007).
[44]J. Neubauer, P. Mergell, U. Eysholdt, and H. Herzel, "Spatio-temporal analysis of irregular vocal fold oscillations: Biphonation due to desynchronization of spatial modes," J. Acoust. Soc. Am. **110**, 3179 (2001).
[45]H. Hatzikirou, W. T. Fitch, and H. Herzel, "Voice instabilities due to source-tract interactions," Acta Acust. **92**, 468 (2006).
[46]S. Kiritani, H. Hirose, and H. Imagawa, "High-speed digital imaging analysis of vocal fold vibration in diplophonia," Speech Commun. **13**, 23 (1993).
[47]T. Wittenberg, M. Tigges, P. Mergell, and U. Eysholdt, "Functional imaging of vocal fold vibration: Digital multi-slice high-speed kymography," J. Voice **14**, 422 (2000).