Computation of physiological human vocal fold parameters by mathematical optimization of a biomechanical model

Anxiong Yang^{a)}

Department of Phoniatics and Pediatric Audiology, University Hospital Erlangen, Medical School, Bohlenplatz 21, 91054 Erlangen, Germany

Michael Stingl

Department of Applied Mathematics II, University of Erlangen-Nuremberg, Martensstrasse 3, 91058 Erlangen, Germany

David A. Berry

The Laryngeal Dynamics Laboratory Division of Head and Neck Surgery, David Geffen School of Medicine at UCLA, 31-24 Rehab Center, 1000 Veteran Avenue, Los Angeles, California 90095-1794

Jörg Lohscheller

Department of Medical Computer Science, Trier University of Applied Sciences, Schneidershof, 54293 Trier, Germany

Daniel Voigt

Department of Linguistics, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany

Ulrich Eysholdt and Michael Döllinger

Department of Phoniatics and Pediatric Audiology, University Hospital Erlangen, Medical School, Bohlenplatz 21, 91054 Erlangen, Germany

(Received 2 July 2010; revised 1 June 2011; accepted 3 June 2011)

With the use of an endoscopic, high-speed camera, vocal fold dynamics may be observed clinically during phonation. However, observation and subjective judgment alone may be insufficient for clinical diagnosis and documentation of improved vocal function, especially when the laryngeal disease lacks any clear morphological presentation. In this study, biomechanical parameters of the vocal folds are computed by adjusting the corresponding parameters of a three-dimensional model until the dynamics of both systems are similar. First, a mathematical optimization method is presented. Next, model parameters (such as pressure, tension and masses) are adjusted to reproduce vocal fold dynamics, and the deduced parameters are physiologically interpreted. Various combinations of global and local optimization techniques are attempted. Evaluation of the optimization procedure is performed using 50 synthetically generated data sets. The results show sufficient reliability, including 0.07 normalized error, 96% correlation, and 91% accuracy. The technique is also demonstrated on data from human hemilarynx experiments, in which a low normalized error (0.16) and high correlation (84%) values were achieved. In the future, this technique may be applied to clinical high-speed images, yielding objective measures with which to document improved vocal function of patients with voice disorders. © 2011 Acoustical Society of America. [DOI: 10.1121/1.3605551]

PACS number(s): 43.70.Aj, 43.70.Bk, 43.70.Gr, 43.72.Ar [CHS]

Pages: 948–964

I. INTRODUCTION

In recent years, voice scientists have utilized high-speed and multiple-camera imaging techniques to intensify their study of vocal fold dynamics. One important part of normal vocal fold vibration is the traveling wave which propagates along the vocal fold mucosa, commonly referred to as “mucosal wave propagation” presented by Hirano¹ and Baer,² and further developed by recent studies (e.g., Refs. 3 and 4). The wave propagates not only in the transverse plane (lateral-longitudinal) but also in the sagittal plane (longitudi-

nal-vertical) and coronal plane (lateral-vertical). The medial edge of the vocal fold exhibits wider and faster movements than more ventral and dorsal regions. Also, larger amplitudes of vibration are observed superiorly along the medial surface of the folds, as compared to more inferior regions.^{3,5}

Healthy vocal folds vibrate more symmetrically and periodically than vocal folds with pathologies.⁶ In clinical diagnosis, morphological pathologies of the larynx such as unilateral vocal fold polyps can be visually observed using a standard endoscope. However, some pathologies may not exhibit static, morphologic alterations, but only manifest themselves dynamically during phonation, e.g., functional dysphonias.⁷ Clinically, such voice disorders may exhibit left-right asymmetries or non-periodic oscillations, and usually result in a

^{a)}Author to whom correspondence should be addressed. Electronic mail: anxiong.yang@uk-erlangen.de

hoarse voice.⁸ These asymmetric and irregular vocal fold vibrations may be induced by pathological changes in biomechanical properties of vocal fold tissues.^{6–12}

To determine whether an acoustic voice signal is abnormal, it is evaluated with respect to its perceived vocal quality, variability, pitch, and loudness.¹³ Such parameters may be dependent on subglottal pressure, vocal fold posture, tension,^{11,14} and morphology.¹⁵ Common and widespread methods for diagnosing laryngeal diseases are generally classified into subjective and objective categories. Trained listeners use running speech for auditory assessment of voice quality, which is a subjective method capturing only the symptomatic information of a laryngeal disease.¹⁶ Indeed, discrepancies between auditory and laryngoscopic findings may occur.¹⁷ Moreover, the diagnosis and differentiation of voice disorders is often marked by ambiguity.⁹ Conventional visual methods like stroboscopy (high spatial resolution) and videokymography [high temporal resolution: approx. 8000 frames per second (fps)] may assist with clinical diagnosis. However, stroboscopy is limited by low temporal resolution,¹⁸ and videokymography is limited by low spatial resolution (i.e., a single line of the sensor data).¹⁹ Not surprisingly, the increased temporal and spatial resolution offered by high-speed (HS) imaging of the vocal folds during phonation may further enhance clinical diagnosis.^{7,20} Clinically, this is performed by using an endoscopic HS digital imaging technique (2000–4000 fps).^{21–23}

Currently, analysis of HS recordings serves as basis for objective and quantitative measures of normal and pathological vocal fold vibrations.^{21,23} However, so far endoscopic *in vivo* HS recordings only permit analysis of horizontal vocal fold dynamics along the vocal fold edge.²¹ While a recent attempt has been made to analyze vertical dynamics,²¹ most of the biomechanical characteristics of three dimensional (3D) mucosal wave propagation during phonation (i.e., medio-lateral, inferior-superior, and anterior-posterior)^{4,5} are not captured by endoscopic imaging.³ Such 3D characteristics were first detected and investigated in *in vitro* laryngeal experiments.^{4,24,25} For deeper analysis of such vibration patterns, a biomechanical numerical 3D-multi-mass-model (3DM)²⁶ was proposed. It was designed to simulate not only lateral and longitudinal vocal fold vibrations, but also vibrations in the vertical direction. The 3DM was a vertical extension of the previous two dimensional multi-mass-model (2DM),²⁰ and was introduced in detail in Ref. 26. Numerical modeling of vocal fold dynamics and physiological interpretation of the biomechanical parameter values may help to increase our understanding of both normal and pathological voice production.²⁷ For example, all the measured aspects of vocal fold dynamics (including amplitudes of vibration, phase delays, velocities and accelerations across a large region of tissue locations) were evaluated to reveal viscoelastic properties of vocal fold tissues.²⁶ Significantly, even biomechanical parameters of pathological vocal folds may be deduced using the 3DM. Thus, the derived parameters may serve as an objective measure to quantify the physiological states of vocal fold vibration. For example, observed left-right asymmetries can be detected and further analyzed by comparing the resultant bilateral, biomechanical parameters of the model. Therefore, the 3DM provides a convenient

way to describe the biomechanical properties of measured 3D vocal fold dynamics.

In recent attempts using the two-mass-models (2MM)^{7,28,29} and the 2DM^{20,30} only 1D trajectories at one or more specific locations on the vocal fold edges along the longitudinal direction were fitted. Classification schemes between healthy and pathological vocal fold vibrations were presented. However, such models considered the vocal fold oscillations only in the lateral dimension. The corresponding dynamic behaviors in the remaining dimensions were negligible, due to undue restriction of the models. In this work, the 3D modeling approach is proposed to be able to reproduce the 3D vibrations of the entire medial surface of the vocal fold. This can be regarded as an important step for linking with the classification of the various 3D vocal fold vibratory modes as well as the production of the acoustic voice signal in future studies.

In order to refine the quantification of the biomechanical properties of human vocal folds and to support clinical diagnosis, the goal is to adapt the 3DM for both *in vitro*^{4,31–33} and *in vivo* 3D HS recordings. To this end, an optimization procedure with a corresponding objective function for the 3DM adaption is presented. Modifications of the 3DM can be imposed by expressing the model parameters in terms of their initial values by introducing optimization factors, which influence the constants of spring, mass, subglottal pressure and rest position, respectively. As objective functions, minimizing the normalized error between the model adapted 3D trajectories and the trajectories to be fitted are chosen. The non-convex objective function contains a large number of local minima which make the optimization procedure complicated.⁷ Hence, a combination of global and local optimization techniques is applied, e.g., Particle Swarm Optimization,³⁴ simulated annealing³⁵ and Powell's direction set method.³⁶ In order to yield appropriate results progressive optimization sub-procedures are implemented along each side, each coronal cross section and each transverse plane within the 3DM. The optimization strategy is focused on a refinement, wherein gradually increasing numbers of individual mass elements are taken into account. Due to characteristics of the 3D vocal fold motions, special weighting coefficients for different dimensions, cross-sections and planes are considered. Overall the methodology of optimization gets more important, since the complexity of the 3D modeling increases. The optimization procedure is verified on 50 synthetically generated vocal fold vibrations encompassing five well-known glottal closure types.^{37–39} Finally, as an essential basis of the application of the 3D modeling to different laryngeal data (excised or *in vivo*) in future, the optimization procedure is demonstrated on data from an *in vitro* experiment. Once the biomechanical parameters have been derived using this method, an interpretable physiological representation of the vocal fold dynamics will be presented and discussed.

II. METHODS

A. 3D biomechanical model

For modeling realistic 3D vocal fold dynamics a 3D-multi-mass-model was previously proposed.²⁶ It consists of

five transverse planes (from inferior to superior) with five coronal cross-sections (from dorsal to ventral). The subglottal pressure P^{sub} acts as a driving force for air flowing in the inferior to superior direction, causing vocal fold oscillations. Figure 1 shows a 3D view of the 3DM.

The model parameters serve as an approximation of tissue properties of the vocal folds, i.e., distributions of masses $m_{i,s}$, dampings $r_{i,s}$ and different spring stiffness coefficients $k_{i,s}^a$ (anchor spring), $k_{i,s}^v$ (vertical spring), $k_{i,s}^l$ (longitudinal spring). Index $s = 1, \dots, 5$ numbers the transverse planes from inferior to superior. Index $i = 1, \dots, 5$ is for labels of mass elements on the right side from dorsal to ventral, and $i = 6, \dots, 10$ is on the left side, see Fig. 1.

The biomechanical model can be described by a system of 50 Ordinary Differential Equations (ODE). For each mass element $m_{i,s}$ the following differential equation holds:

$$m_{i,s}\ddot{x}_{i,s} = \vec{F}_{i,s}^A + \vec{F}_{i,s}^V + \vec{F}_{i,s}^L + \vec{F}_{i,s}^C + \vec{F}_{i,s}^D. \quad (1)$$

The 3D position of each mass element is denoted as $x_{i,s} = [x_{i,s}, y_{i,s}, z_{i,s}]^T$. Its first and second derivatives with respect to time are the velocity $\dot{x}_{i,s}$ and acceleration $\ddot{x}_{i,s}$, respectively. The applied forces acting on each mass element are defined as follows.

- (1) The anchor force $\vec{F}_{i,s}^A$ is due to the anchor spring $k_{i,s}^a$ and the damping $r_{i,s}^a$.
- (2) The vertical coupling force $\vec{F}_{i,s}^V$ arises from the vertical spring stiffness $k_{i,s}^v$ and the damping coefficient $r_{i,s}^v$.
- (3) The longitudinal coupling force $\vec{F}_{i,s}^L$ is caused by the longitudinal spring stiffness $k_{i,s}^l$ and the corresponding damping coefficient $r_{i,s}^l$.
- (4) The collision impact force $\vec{F}_{i,s}^C$ is derived from the collision restoring spring $k_{i,s}^c$.

FIG. 1. Schematic representation of the 3DM for biomechanical modeling of human vocal fold dynamics. Every mass element is elastically connected to a rigid body with an anchor spring. Moreover, at each side every mass is connected to its adjacent masses with springs in vertical and longitudinal directions. The indices denote the different planes s and columns i for the mass elements.

- (5) The driving force $\vec{F}_{i,s}^D$ is generated by the glottal flow. It is derived from the subglottal pressure P^{sub} and the corresponding effective area using Bernoulli's Law.

The different damping coefficients $r_{i,s}^a, r_{i,s}^l, r_{i,s}^v$ are dependent on $m_{i,s}$ and $k_{i,s}^a$ of the adjacent mass elements. The collision impact force $\vec{F}_{i,s}^C$ plays a vital role in phonation, especially for higher resonant frequency components.⁴⁰ With the aid of the 3DM, not only the small amplitude oscillation but also the large amplitude oscillation are modeled. Nonlinearities of the deflection of muscles and ligaments cannot be ignored at large amplitudes of vibration and at high subglottal pressure.^{41,42} As underlined by Titze⁴³ human tissue does not exhibit linear stiffness characteristics. Also, Titze and Durham⁴⁴ demonstrated that the fundamental frequency is significantly affected by nonlinearity in the stiffness. Therefore, to incorporate variation of fundamental frequency of large/small oscillations with subglottal pressure into the 3DM, the nonlinear coefficient η for spring stiffness is set to 100 cm^{-2} .^{41,45}

The detailed mathematical definition of the model and the forces were described in Ref. 26. The ODE-system is solved by using the Runge-Kutta algorithm with a step-size of 0.25 ms.⁷

B. Automatic parameter optimization

In order to match the model dynamics with the dynamics of human vocal folds, the model parameters must be optimized. Compared to previous studies applying 2MM⁴¹ and 2DM,²⁰ the 3DM has a high dimensionality captured by the spring stiffnesses ($50k_{i,s}^a, 60k_{i,s}^l, 60k_{i,s}^v$), masses ($50m_{i,s}$), damping coefficients ($50r_{i,s}^a, 60r_{i,s}^l, 60r_{i,s}^v$), rest positions ($50x_{i,s}^r$ in 3D), subglottal pressure P^{sub} , and boundary conditions (including the vertical anchors and fixed positions at the ventral/dorsal ends of vocal folds). Overall, the 3DM yields 531 degrees of freedom.

1. Initialization of model and optimization parameters

First, a default model configuration is set. That is, standard model parameters (e.g., stiffness, mass, thickness, rest position, etc.) are chosen to serve as an initial approximation of biomechanical properties of the vocal fold.^{26,46} The initial values for the model parameters are chosen not only on the basis of previous 2MM⁴¹ and 2DM^{20,30} configurations, but also on the basis of propagation properties of the mucosal wave, as observed in human laryngeal experiments.^{4,25} A detailed description of the applied standard model parameters was presented previously.²⁶ Thus, the configuration of the 3DM is imposed by expressing the model parameters $P^{sub}, m_{i,s}, k_{i,s}, x_{i,s}^r$ in terms of the previously introduced standard model parameters $\tilde{P}^{sub}, \tilde{m}_{i,s}, \tilde{k}_{i,s}, \tilde{x}_{i,s}^r$. In order to reduce the dimensionality of the parameter space and to decrease computational costs, certain model parameters were coupled together by introducing an optimization parameter set:³⁰ $\mathbf{Q} := (Q_p, Q_{i,s}, Q_r)$, with index i,s for each mass element $m_{i,s}$ within the 3DM. The parameters $Q_{i,s} := (Q_{i,s}^a, Q_{i,s}^l, Q_{i,s}^v)$

influence the stiffness of anchor spring $k_{i,s}^a$, the stiffness of longitudinal spring $k_{i,s}^l$ and vertical spring $k_{i,s}^v$, respectively. Using the optimization parameters $Q_{i,s}^a$, modifications of mass and anchor spring parameters (i.e., lateral stiffness) can be performed simultaneously, due to the mass-spring systems as well as the highest vibratory displacements in the lateral direction.^{4,26} In addition, Q_p affects the subglottal pressure P^{sub} , and $Q_r := (Q_r^{i,s,x}, Q_r^{i,s,y}, Q_r^{i,s,z})$ modifies the rest position $\mathbf{x}_{i,s}^r$ of each mass element. Overall, the various model parameters can be derived completely from the initial standard model parameters:^{7,20}

$$P^{sub} = \tilde{P}^{sub} Q_p, \quad (2)$$

$$m_{i,s} = \tilde{m}_{i,s} / Q_{i,s}^a. \quad (3)$$

$$k_{i,s}^\xi = \tilde{k}_{i,s}^\xi \cdot Q_{i,s}^\xi, \quad \xi \in \{a, l, v\}, \quad (4)$$

$$\mathbf{x}_{i,s}^r = [\tilde{x}_{i,s}^r \cdot Q_r^{i,s,x}, \quad \tilde{y}_{i,s}^r \cdot Q_r^{i,s,y}, \quad \tilde{z}_{i,s}^r \cdot Q_r^{i,s,z}]^T. \quad (5)$$

Moreover, it is also considered that the damping coefficients $r_{i,s}$ are modified as $\tilde{r}_{i,s} \cdot Q_d$. Corresponding scaling factor $\tilde{Q}_d$ is initialized to 1, since the damping coefficients were initialized with the relationship among the standard values of stiffness and mass as well as damping ratio.²⁶ However, they are negligible in validation of the optimization, because their effects on model-generated dynamics are small enough, compared to the stiffness and mass.²⁶ After reducing the dimensionality, it results in 491 Q scaling factors to be optimized.

Since the model contains 50 simple mass-spring-oscillators with mass $m_{i,s}$ and reciprocal coupling spring constant $k_{i,s}$, the fundamental frequency $f_{i,s}$ for each mass can be approximated as follows:⁷

$$f_{i,s} = \frac{1}{2\pi} \sqrt{\frac{k_{i,s}}{m_{i,s}}}. \quad (6)$$

To get initial parameters $\tilde{Q}_{i,s}$ Eq. (6) can be rewritten:⁷

$$\tilde{Q}_{i,s}^\xi = 2\pi f_{i,s} \sqrt{\frac{\tilde{m}_{i,s}}{\tilde{k}_{i,s}^\xi}}, \quad (7)$$

which shows that the optimization parameter $Q_{i,s}$ primarily affects the fundamental frequency of oscillation.²⁸

The optimization parameter Q_p mainly influences the oscillation amplitude.³⁰ An appropriate value of the initial

optimization parameter $\tilde{Q}_p$ yields smaller amplitude differences between the curves $\mathbf{c}_{i,s}[n]$ to be fitted and the model generated curves $\mathbf{c}_{i,s}^M[n]$, where $n = 1, \dots, N$ denotes the frame number. Namely, by comparing the amplitude differences under different values (2–35 cm H₂O) of subglottal pressure p , the initial value of $\tilde{Q}_p$ is determined:

$$\tilde{Q}_p := \frac{1}{\tilde{P}^{sub}} \cdot \arg \min_p \left(\frac{1}{5 \cdot 10} \sum_{s=1}^5 \sum_{i=1}^{10} \|\mathbf{A}_{i,s} - \mathbf{A}_{i,s}^M(p)\|_2 \right), \quad (8)$$

$$\mathbf{A}_{i,s} := \frac{1}{2} \cdot \left\| \max_n (\mathbf{c}_{i,s}[n]) - \min_n (\mathbf{c}_{i,s}[n]) \right\|_2, \quad (9)$$

$$\mathbf{A}_{i,s}^M(p) := \frac{1}{2} \cdot \left\| \max_n (\mathbf{c}_{i,s}^M[n]) - \min_n (\mathbf{c}_{i,s}^M[n]) \right\|_2, \quad (10)$$

where $\mathbf{A}_{i,s}$ denotes the amplitude of the experimental 3D data. $\mathbf{A}_{i,s}^M(p)$ is the amplitude of the 3D curves generated by 3DM with a specified subglottal pressure value.

Additionally, the optimization parameter $\tilde{Q}_r$ is initialized to 1, since the rest positions were initialized with the standard values.

2. Objective function

In general, the optimization of the 3DM compares the dynamic properties between the adapted model-generated 3D trajectories and the experimental 3D trajectories, such as amplitudes and corresponding correlations.

In this work, unlike the definition of the objective function in 2DM²⁰ which also used a glottis area consistency measure, a trajectory consistency measure in 3D is applied. Additionally, according to the previous studies of laboratory larynx experiments,^{4,24,47} it is revealed that the displacements along three orthogonal directions are highest in the superior region of the vocal folds. Moreover, the vocal fold displacements in the lateral direction are significantly higher than those in the other directions, while the longitudinal displacements are the smallest. In consideration of such characteristics, different weighting coefficients for each mass element in each dimension, each coronal cross section and each transverse layer are taken into account within the computation of the objective function, so that the main components of vocal fold dynamics can be determined. For the purpose of minimizing the error between the experimental 3D trajectories $\mathbf{c}_{i,s}[n]$ and the theoretical trajectories $\mathbf{c}_{i,s}^M[n]$ of the 3DM, the following objective function r which measures the error is suggested:

$$\Gamma(Q) := \sum_{a=1}^5 \sum_{i=1}^{10} \left\{ g_i^m \cdot g_s^p \cdot \left[\frac{\sum_{n=1}^N \|g_\mu^d \cdot (\mathbf{c}_{i,s}[n] - \mathbf{c}_{i,s}^M[n])\|_2}{\sum_{n=1}^N \|g_\mu^d \cdot \mathbf{c}_{i,s}[n]\|_2} \right] \right\}, \quad (11)$$

with

$$\begin{aligned} \mathbf{c}_{i,s} &= [c_{i,s,x}, c_{i,s,y}, c_{i,s,z}]^T, \\ \mathbf{c}_{i,s}^M &= [c_{i,s,x}^M, c_{i,s,y}^M, c_{i,s,z}^M]^T, \end{aligned} \quad (12)$$

where μ denotes x- (lateral), y- (longitudinal), and z-dimension (vertical), respectively. $i = 1, \dots, 10$ denotes the number of the mass elements at each plane. $s = 1, \dots, 5$ denotes the number of the planes. Equation (11) corresponds to the energy of the error normalized to the energy of the experimental trajectories.³⁰ Hence, the objective function Γ can be regarded as a measure for the normalized relative error.

The weighting coefficients for 3D trajectories are introduced as g_μ^d , g_i^M and g_s^p . They refer to different dimensions (superscript d) and different mass elements (superscript m) as well as different planes (superscript p), respectively. All of the weighting coefficients are located in the range between 0 and 1. By observing the mucosal wave dynamics of the vocal folds in hemilarynx experiments^{3,4,24} and excised canine larynx experiments,^{47,48} the lateral displacement (x-component) is in general highest. The corresponding local vertical displacement maxima (z-component) predominantly range between longitudinal (y-component) and lateral values.^{4,25} Based on these reasons, the weighting g_μ^d is derived as

$$g_\mu^d := \frac{\sum_{s=1}^5 \sum_{i=1}^{10} [\max_n(c_{i,s,\mu}[n]) - \min_n(c_{i,s,\mu}[n])]}{\sum_{s=1}^5 \sum_{i=1}^{10} \left\| \max_n(c_{i,s}[n]) - \min_n(c_{i,s}[n]) \right\|_1}, \quad (13)$$

where $n = 1, \dots, N$.

In addition, the 3D displacements differ along the length of the vocal folds. Usually the highest amplitude of displacement is found midway between anterior and posterior ends of the fold.^{4,49} The 3D displacements at the anterior and posterior ends of the vocal folds are lower, being almost fixed during oscillation.^{24,50} Larger and faster movements occur in the central section of the vocal fold than near the glottal extremities.⁵ According to these properties, we assume the weighting g_i^m gm corresponding to each mass element as follows:

$$g_i^m := \frac{\sum_{s=1}^5 \left(\max_{n,n'} \|\mathbf{c}_{i,s}[n] - \mathbf{c}_{i,s}[n']\|_2 \right)}{\sum_{i=1}^{10} \sum_{s=1}^5 \left(\max_{n,n'} \|\mathbf{c}_{i,s}[n] - \mathbf{c}_{i,s}[n']\|_2 \right)}, \quad (14)$$

where $n' = n + 1$ denotes the frame number. Here, $n = 1, \dots, N - 1$.

From earlier studies of hemilarynx experiments, it was concluded that the superior part of the vocal fold exhibits greater amplitudes of vibration.⁴⁷ Therefore, for optimization, 3D movements in the upper planes receive relatively more attention than movements in the lower planes. The weighting g_s^p for each plane from inferior to superior is defined as follows:

$$g_s^p := \frac{\sum_{i=1}^{10} \left(\max_{n,n'} \|\mathbf{c}_{i,s}[n] - \mathbf{c}_{i,s}[n']\|_2 \right)}{\sum_{s=1}^5 \sum_{i=1}^{10} \left(\max_{n,n'} \|\mathbf{c}_{i,s}[n] - \mathbf{c}_{i,s}[n']\|_2 \right)}, \quad (15)$$

where $n = 1, \dots, N - 1$.

The weighting coefficients are computed based on 3D trajectories from *in vitro* hemilarynx experiments:⁴

$$g_\mu^d = (0.5, 0.1, 0.4), \quad g_s^p = (0.1, 0.1, 0.3, 0.3, 0.2), \quad (16)$$

$$g_{i=1,\dots,5}^m = g_{i=6,\dots,10}^m = (0.095, 0.1, 0.11, 0.1, 0.095). \quad (17)$$

These values are used in the following validation and application of the optimization procedure.

A second criterion for judging the quality of fit is the correlation coefficient κ . It measures how closely the shapes of the optimized 3D trajectories match the experimental 3D trajectories. Thus, the similarity and phases of the trajectories are taken into account:

$$\kappa := \sum_{s=1}^5 \sum_{i=1}^{10} \left[g_i^m \cdot g_s^p \cdot \left(g_x^d \cdot \kappa_{i,s,x} + g_y^d \cdot \kappa_{i,s,y} + g_z^d \cdot \kappa_{i,s,z} \right) \right], \quad (18)$$

with

$$\kappa_{i,s,\mu} := \frac{\langle c_{i,s,\mu}[n], c_{i,s,\mu}^M[n] \rangle}{\|c_{i,s,\mu}[n]\|_2 \|c_{i,s,\mu}^M[n]\|_2} \cdot 100\%. \quad (19)$$

Now, the condition for the quality of the optimization procedure is defined as

$$(\Gamma < 0.2) \wedge (\kappa \geq 75\%). \quad (20)$$

If this is fulfilled, the optimization quality can be regarded as sufficiently successful. The lower bound 75% for κ is on the basis of results of Wurzbacher *et al.*²⁸ The upper bound 0.2 for Γ is defined in accordance with the relative endoscopic HS image processing error.²⁸

3. Optimization algorithm

Since the resulting optimization problem is nonconvex,⁷ applying only local (gradient) optimization techniques will not yield sufficiently accurate results. Therefore, various combinations of global and local optimization algorithms are applied. During the optimization procedure, the global optimization algorithm is used to determine preliminary parameter values. Then, on the basis of the preliminary minimum, a local optimization algorithm is applied to find parameter values that are even closer to the global minimum of the objective function Γ . However, reaching the global minimum cannot be guaranteed in nonconvex optimization problems. For controlling the switch of the global optimizations, the following heuristic condition is defined as $\kappa < 80\%$. If this condition is fulfilled, both the global and local optimization

algorithms are sequentially applied. Otherwise, only the local optimization algorithm is carried out. This switch condition is implemented once at the beginning of each phase within each optimization sub-procedure, which is introduced in the next section.

For the selection of the optimization algorithm, conjugate gradient algorithms (e.g., Fletcher-Reeves) and variable metric algorithms (e.g., Broyden-Fletcher-Goldfarb-Shanno)³⁶ have proven to be inadequate and relatively time-consuming because of the complex partial derivatives of the objective function. Thus, in this work, the more suitable global methods for minimizing the objective function are the Particle Swarm Optimization (PSO) and the Simulated Annealing (SA). Briefly, the PSO is a population-based stochastic optimization technique.^{34,51} It was first used to model the social behavior of bird flocking, bee swarming or fish schooling.⁵² Similar to Genetic Algorithms (GA),⁵³ PSO also uses many evolutionary computation techniques,⁵⁴ including random searches, fitness values, iteration time, and so on. The system is initialized with a population of potential solutions and searches for optima by updating generations. Each individual (particle) traverses the problem space with its own vector, which determines its next trajectory following the optimum particle. However, no evolution operators such as crossover and mutation are applied in PSO. Therefore, PSO has the advantage of simply controlling and robustly adjusting parameters. The SA is known to work as a generic probabilistic metaheuristic for optimization problems of large scale, especially the ones in which a desired global extremum is hidden among many local extrema. A detailed description of Simulated Annealing can be found in Ingber.³⁵ Accordingly, we choose a combination between PSO and SA to globally seek the minimum of the objective function, since in a study of metaheuristic methods⁵⁴ Ercan confirmed that PSO-SA combinations outperform the basic PSO algorithm. A heuristic condition $(\Gamma \geq 0.1) \vee (\kappa \leq 90\%)$ is defined as a switch to the PSO-SA combination. Namely, if it is fulfilled, the PSO will be implemented first, followed by the SA (with the results of the PSO used as initial solutions for the SA). Otherwise, only the SA will be executed. Moreover, for the local optimization algorithm, the Powell's Direction Set Method (PDSM)³⁰ [which belongs to the class of conjugate direction algorithms] is adapted, which in this instance has been proven to be faster and more stable than other optimization algorithms such as the Nelder-Mead algorithm.⁷

Compared to the use of any single algorithm (PSO, SA, or PDSM), the consecutive use of different optimization algorithms improves the consistency and stability in approximating the global minimum. However, it may also pose a computational burden. To better clarify the proposed optimization algorithm, a flow chart is sketched in Fig. 2. It is implemented in each phase of the optimization procedure, which is introduced in the following section.

C. Progressive optimization procedure

In optimizing the 3DM, a progressive concept is introduced, wherein the search space consists of the following

FIG. 2. Flow chart of the combined optimization algorithms. PSO, SA, and PDSM are applied to optimize the parameters Q to fit the model-generated dynamics $c_{i,s}^M[n]$ to the vocal fold vibrations $c_{i,s}[n]$. This is performed in each optimization phase, Fig. 3.

optimization parameters: $(Q_p, Q_{i,s}, Q_r)$. Due to the location of each mass element in 3D, three aspects (plane, side, cross section) are taken into account, which affect the optimization of the 3DM. In accordance with various combinations among the three aspects, the so-called progressive optimization procedure consists of three main coarse optimization sub-procedures (Fig. 3). The basic idea of the optimization procedure is to gradually adapt the model parameters from a rough state to a more refined state, to overcome the difficulties of multidimensional optimization. In this work, the simple global optimization approach has proven to be inconsistent in fitting model-generated dynamics to experimental data. However, using a hierarchical coarse optimization, the global minimum can be consistently and stably approached.

Within each sub-procedure, three phases are implemented one after another (see Fig. 4). Moreover, within each phase, the combination (Sec. II B 3) between global optimization algorithm (PSO-SA) and local optimization algorithm (PDSM) is implemented. Switching from the current phase to the next phase occurs when the objective function r falls below a predefined threshold or when the maximum iteration value is reached. The configuration of each individual optimization procedure is detailed below:

- (1) PSO. The number of particles is set to 100. The maximum number of iterations is 1000. An inertia weight of the particle is 0.73. Two acceleration constants of the particle, which represent the cognitive and social learning of the particle, are set to 1.49.
- (2) SA. The number of steps is set to 2. The maximum iteration number at each step is 300. The temperature is

FIG. 3. General flow chart of the hybrid optimization procedure. $(\hat{Q}_p, \hat{Q}_{i,s}, \hat{Q}_r)$ represent the initial optimization parameters. $(\hat{Q}_p, \hat{Q}_{i,s}, \hat{Q}_r)$ represent the best optimization parameters. There are three *coarse* sub-procedures and a fine optimization process. Each *coarse* sub-procedure has three *phases*. In this flow chart, m, n, l temporarily denote the index of the *coarse* sub-procedure, the *phase*, and the loop, respectively.

initialized to 10. The density of the randomly chosen points of the initial simplex is 0.1. For determining a local extremum, a so-called fractional tolerance is set to 0.001, which measures the difference between the objective function values in two consecutive steps. The non-linear constant is set to 2. The more nonlinear the objective function is, the lower the non-linear constant should be set.

- (3) PDSM. The maximum iteration number is set to 600. A parabolic extrapolation algorithm “Parabolic-Bracketing” is used as the bracketing algorithm. To determine where an extremum value lies, a minimum distance between the end points of any bracketing interval is set to 0.5. The localization algorithm in PDSM is Brent’s algorithm.³⁶ Similarly, a fractional tolerance is set to 0.001. In addition, a tolerance value which is used to check the exit condition for PDSM is set to 0.5. As a rule, the smaller

the tolerance value, the more accurate the result of PDSM algorithm. However, the algorithm also becomes more computationally demanding.

The value assignments are based on heuristic analysis and experience. More detailed descriptions of the applied algorithms can be found in Refs. 34 and 36.

Additionally, in each phase, control parameters $\alpha_\nu := (\alpha_\nu^a, \alpha_\nu^l, \alpha_\nu^v, \alpha_\nu^r)$ serve as scale factors to control the modification of the optimization parameters $\mathbf{Q}$:

$$\mathbf{Q}_{\text{new}} = \alpha_\nu \cdot \mathbf{Q}_{\text{old}}, \quad (21)$$

where $\mathbf{Q}_{\text{old}}$ and $\mathbf{Q}_{\text{new}}$ are the optimization parameters before and after each phase, respectively. Thus, the optimization parameters can be continuously updated in each phase.

FIG. 4. Block charts for the three course optimization sub-procedures. Each sub-procedure is divided into three phases: Along each row, the intercommunities are vertical, lateral, and longitudinal, respectively.

The concept of three sub-procedures and their individual conditions for switching to the next sub-procedure are introduced in the following sections. Each sub-procedure (coarse 1–3) is terminated if its cycle index reaches a recommended limit value of $l = 10$.

In addition, at the end of the optimization procedure a fine optimization process (Fig. 3) is implemented. In this process, all of the model parameters are one-to-one optimized, in order to further improve the optimization results.

Overall, in each phase of the optimization procedure, the objective function r either decreases or remains momentarily unchanged. Γ is bounded by greatest lower bound 0, which indicates no error.

1. Sub-procedure coarse 1

The purpose of this sub-procedure is to roughly adapt the model dynamics to the experimental data. It is continuously performed until the following condition is no longer fulfilled:

$$(\Gamma_l < \Gamma_{l-1}) \wedge (k < 92\%) \wedge (l < 10), \quad (22)$$

where the subscript l denotes the index of the loop. $l - 1$ and l represent two consecutive loops. The threshold value of κ is defined as 92%, based on experience gained in former work.²⁸ In general, it is separated into three phases [Figs. 4(a)–4(c)]:

- (1) Vertical optimization. Parameters for each plane are modified with the same factor $\alpha_{\nu=1,\dots,5}$.
- (2) Lateral optimization. Parameters for each side are modified with the same factor α_{ν} with

$$\nu = \begin{cases} 1 & \text{for } i = 1, \dots, 5, \\ 2 & \text{for } i = 6, \dots, 10. \end{cases} \quad (23)$$

- (3) Longitudinal optimization. Parameters for each cross section are modified with the factor α_{ν} with

$$\nu = \begin{cases} i & \text{for } i = 1, \dots, 5, \\ i - 5 & \text{for } i = 6, \dots, 10. \end{cases} \quad (24)$$

2. Sub-procedure coarse 2

The parameters resulting from coarse 1 are set as initial values for this sub-procedure. The exit-condition for coarse 2 is defined as follows:

$$(\Gamma_l < \Gamma_{l-1}) \wedge (\Gamma > 0.04) \wedge (l < 10), \quad (25)$$

where 0.04 is a heuristic threshold value for the normalized error, in accordance with results presented in Wurzbacher *et al.*²⁸ Three phases are employed as follows [see Figs. 4(d)–4(f)]:

- (1) Vertical-lateral optimization Parameters for each side of each plane are modified with the same factor α_{ν} , with:

$$\nu = \begin{cases} s, & \text{for } i = 1, \dots, 5, \\ s + 5, & \text{for } i = 6, \dots, 10. \end{cases} \quad (26)$$

- (2) Lateral-longitudinal optimization Parameters for each side of each cross section are modified with the same factor α_{ν} , with $\nu = i$.
- (3) Longitudinal-vertical optimization Parameters for each pair of mass-spring-oscillators are modified with the same factor α_{ν} with:

$$\nu = \begin{cases} 5 \cdot (s - 1) + i, & \text{for } i = 1, \dots, 5, \\ 5 \cdot (s - 2) + i, & \text{for } i = 6, \dots, 10. \end{cases} \quad (27)$$

3. Sub-procedure coarse 3

The aim of this sub-procedure is to adapt the optimization parameters even more thoroughly, on the basis of the preceding coarse 2. The exit-condition for coarse 3 is defined as follows:

$$(\Gamma_l < \Gamma_{l-1}) \wedge (k < 95\%) \wedge (l < 10), \quad (28)$$

where 95% is defined as the threshold value of κ in this sub-procedure, based on former work²⁸ as well as the preceding sub-procedures. Three phases are implemented as follows [see Figs. 4(g)–4(i)]:

- (1) Plane optimization. Parameters for the 10 mass elements ($m_{1,s}, \dots, m_{10,s}$) at the current plane s are allowed to be modified with different control parameters $a_{\nu=1,\dots,10}$,

while the parameters for mass elements at the other four planes are not adjusted.

- (2) Side optimization. Parameters for 25 mass elements on the current side are modified with different control parameters a_{ν} with $\nu = 1, \dots, 25$.
- (3) Cross section optimization. Parameters for 10 mass elements at the current cross-section are modified with different control parameters $a_{\nu=1,\dots,10}$, while the parameters for the mass elements at the other cross-sections are not modified.

D. Validation and application

1. Synthetic data

In order to evaluate the reliability and capability of the optimization procedure, synthetically generated data are analyzed. Fifty sets of synthetic, laterally-symmetric 3D trajectories are produced by the 3DM with known optimization parameter values (see Table I). To simulate a wide variety of vocal fold vibrations (i.e., frequencies for males and females, amplitudes between 0.3 and 1.2 mm),⁵⁵ the scale factors of the predefined model parameters are randomly set in the following range: $Q_p^S \in \{0.5, 0.6, \dots, 2.7\}$, $Q_{i,s}^S \in \{0.5, 0.6, \dots, 3.1\}$, $Q_r^S \in \{0.6, 0.7, \dots, 1.2\}$. By using these values, masses, elastic modulus of the model, the subglottal pressure, and the rest positions are correspondingly modified.

From a physiological and clinical point of view, the glottal closure type (GCT) is described as an important aspect of laryngeal behavior.^{38,56} Hence, in generating synthetic data, five well-known glottal closure types are considered: rectangle (RA), hourglass (HG), triangular-pointed dorsal (TPD), triangular-pointed ventral (TPV) and convex (CV) (see Fig. 5 and Table I). Since the mean fundamental frequency for males is considerably lower than for females,⁵⁷ the frequency range of the 50 synthetic dynamics is roughly divided into two groups: ≤ 120 Hz for male, ≥ 180 Hz for female. Results from Sulter *et al.*^{58,59} indicate that GCT and glottal chink locations are different for males and females. Thus, in this work, GCT and gender are regarded as two factors that influence optimization accuracy. A variety of vocal fold behaviors was produced for different GCT and gender types, in order to validate optimization of the 3DM as thoroughly as possible.

TABLE I. Fifty synthetic data sets with predefined parameters (Q_p^S , $Q_{i,s}^S$, Q_r^S) to be optimized. The data are separated by glottal closure types (Fig. 5) and gender.

GCT	Gender	No.	Q_p^S	$Q_{i,s}^S$	Q_r^S	Frequency (Hz)
RA	male	5	0.8 ± 0.2	0.9 ± 0.2	0.9 ± 0.2	110 ± 10
	female	5	2.2 ± 0.3	1.5 ± 0.6	0.8 ± 0.2	216 ± 22
HG	male	5	0.7 ± 0.1	0.9 ± 0.2	0.9 ± 0.1	110 ± 11
	female	5	2.1 ± 0.4	1.4 ± 0.5	0.8 ± 0.1	206 ± 19
TPD	male	5	0.7 ± 0.1	0.9 ± 0.1	0.9 ± 0.2	114 ± 6
	female	5	1.9 ± 0.3	1.5 ± 0.6	1.0 ± 0.2	209 ± 15
TPV	male	5	0.9 ± 0.1	0.9 ± 0.2	0.9 ± 0.2	113 ± 10
	female	5	1.8 ± 0.1	1.4 ± 0.5	0.9 ± 0.1	204 ± 17
CV	male	5	0.9 ± 0.3	0.9 ± 0.2	21.0 ± 0.1	114 ± 5
	female	5	2.0 ± 0.6	1.5 ± 0.6	0.8 ± 0.2	226 ± 33

FIG. 5. Schematic representation of occurring glottal closure-types and corresponding modeling: (a) rectangle (RA), (b) hourglass (HG), (c) triangular-pointed dorsal (TPD), (d) triangular-pointed ventral (TPV), (e) convex (CV).

An accuracy value is defined to measure how closely the optimized parameters ($\hat{Q}_p, \hat{Q}_{i,s}, \hat{Q}_r$) approximate the pre-defined values ($Q_p^S, Q_{i,s}^S, Q_r^S$):

$$\lambda := \left(1 - \frac{|\hat{Q} - Q^S|}{Q^S} \right) \cdot 100\%. \quad (29)$$

Overall, the assessments of similarity between the adapted dynamics $\mathbf{c}_{i,s}^M[n]$ and the synthetic dynamics $\mathbf{c}_{i,s}^S[n]$ are implemented by using the following: the accuracy λ of optimization parameters, the correlation coefficient κ verifying the similarity between both 3D trajectories, and the objective function (i.e., normalized error) Γ . In this way, the performance of the optimization procedure and the corresponding reproducibility of the optimization results could be evaluated.

2. Hemilarynx experimental data

As a prelude to the future optimization of both *in vivo* and *in vitro* with physiological relevance,⁴ the optimization procedure is also applied to an *in vitro* hemilarynx experimental data set, which includes recorded 3D trajectories of 25 marker-points attached along the medial vocal fold surface. Vertical distance between marker-points was approximately 1.7 ± 0.2 mm, and horizontal distance was approximately 2.0 ± 0.2 mm. A detailed description of the hemilarynx experiments can be found in Refs. 4, 25, and 31. For comparison with the dynamics of the 3DM (which includes both left and right vocal folds), bilateral symmetry was assumed for the *in vitro* hemilarynx data set, since the dynamics of only one vocal fold could be observed in hemilarynx experiments. The fundamental frequency was 120 Hz, and the subglottal pressure was 22.6 cm H₂O.

With the proposed parameter optimization procedure, a set of optimization parameters ($\hat{Q}_p, \hat{Q}_{i,s}, \hat{Q}_r$) is derived from the adapted 3DM. To adapt the 3D rest positions of the 50 mass elements (i.e., for each mass element in each

dimension) 150 $\hat{Q}_r$ are computed, since the adaptations of the complex marker-point locations would not be accurate if only one scale factor were used to modify all of the 3D rest positions.

III. RESULTS

A. Validation of optimization

1. Global measures

Table II shows the accuracy λ , the correlation coefficient κ and the normalized error Γ for the fifty synthetic data sets. The best value of r occurred in the glottal closure type RA with 0.05 ± 0.03 , while the worst occurred in HG and TPV with 0.09 ± 0.03 . For κ , the best value ($97\% \pm 2\%$) occurred in TPD and CV, while the worst ($94\% \pm 2\%$) was in HG. The value of λ was the highest in RA with $92\% \pm 5\%$, while the lowest ($89\% \pm 3\%$) was in HG. The measures by different genders were essentially equal, except for Γ (better value for males). Overall, the measures (Γ , κ , λ) were 0.07 ± 0.03 , $96\% \pm 2\%$ and $91\% \pm 4\%$, averaged

TABLE II. Optimization results for the synthetic data: The global accuracy λ , correlation coefficient κ and objective function Γ exhibit sufficient good performance. Each glottal closure type covers 10 subjects (5 male, 5 female).

GCT	Γ	κ	λ
RA	0.05 ± 0.03	$96\% \pm 2\%$	$92\% \pm 5\%$
HG	0.09 ± 0.03	$94\% \pm 3\%$	$89\% \pm 3\%$
TPD	0.07 ± 0.03	$97\% \pm 2\%$	$90\% \pm 5\%$
TPV	0.09 ± 0.03	$95\% \pm 3\%$	$92\% \pm 2\%$
CV	0.05 ± 0.02	$97\% \pm 2\%$	$91\% \pm 3\%$
total	0.07 ± 0.03	$96\% \pm 2\%$	$91\% \pm 4\%$
Gender	Γ	κ	λ
male	0.05 ± 0.03	$96\% \pm 2\%$	$91\% \pm 5\%$
female	0.09 ± 0.03	$96\% \pm 3\%$	$91\% \pm 3\%$

over all 50 synthetic data sets. The corresponding ranges for these values were 0.007–0.153, 90%–100%, and 83%–99%, respectively. Additionally, the given symmetry ratio (perfect symmetry = 1) right to left was reproduced as 1.01 ± 0.08 in an acceptable range (0.8 to 1.2).

To illustrate the optimization results, Fig. 6 compares the adapted 3D trajectories $c_{4,4}^M[n]$ and the synthetic 3D trajectories $c_{4,4}^S[n]$, corresponding to the mass element $m_{4,4}$ of a subject with a rectangular glottal closure pattern and a fundamental frequency of 120 Hz.

2. Reproducibility of parameters

To determine the reproducibility of the parameters, the individual accuracies ($\lambda_p, \lambda_{i,s}, \lambda_r$) for the predefined optimization parameters ($Q_p^S, Q_{i,s}^S, Q_r^S$) were derived as: $92 \pm 5\%$, $91\% \pm 9\%$, and $97\% \pm 3\%$, averaged over all of the synthetic subjects. The lowest value of the individual accuracy was for $Q_{i,s}^S$, whereas the highest was for Q_r^S . To facilitate an intuitive feel for the reproducibility of the parameters, the relationship between the adapted optimization parameters

and the predefined optimization parameters is illustrated in Fig. 7. In each chart, the regression lines with their respective regression functions are sketched. The corresponding regression fits were high percentage (i.e., close to one), and their disturbance terms were low. Additionally, the 95% confidence interval is very narrow: the distances between the regression lines and the boundaries of the regions were 0.28, 0.27, and 0.08.

3. Influence from weighting coefficients

In order to examine whether the measures of the optimization results influenced the weighting coefficients, the variations of the individual accuracies $\lambda_{i,s}$ at different planes are shown in Fig. 8. Higher accuracies were achieved at the superior planes $s=3, 4$ with $\lambda_{i,3} = 93\% \pm 6\%$ and $\lambda_{i,4} = 95\% \pm 5\%$ averaged over all mass elements, while the lowest was at the inferior plane $s=1$. Additionally, statistical analysis (analysis of variance) showed that optimization accuracy was statistically significantly dependent on plane ($p < 0.01$).

FIG. 6. (Color online) Exemplary results of the 3D dynamics of the mass element $m_{4,4}$ are given. Synthetic trajectories (solid lines) and optimized trajectories (dotted lines) are presented for all three displacement directions. The corresponding accuracy values are: $(\Gamma, \kappa, \lambda) = (0.05, 97\%, 96\%)$.

FIG. 7. Comparison between the adapted optimization parameters ($\hat{Q}_p$, $\hat{Q}_{i,s}$, $\hat{Q}_r$) and pre-defined values (Q_p^S , $Q_{i,s}^S$, Q_r^S). The solid lines indicate the regression lines for the optimization parameters. The corresponding regression functions are shown including the 95% confidence interval. The distances between the regression lines and the confidence interval boundaries are indicated (Δ_a , Δ_b , Δ_c exhibiting the reliability of the algorithm).

4. Influence from GCT and gender

As shown in Table II, the optimization results may be affected by different GCT and gender. To further investigate how a response is affected by the two factors, two-way analysis of variance (analysis of variance) was applied. It shows that GCT and gender both exhibit statistically significant effects for Γ at the 99% confidence interval (significance level $\alpha=0.01$, probability of a Type I error), since ratios of F-test (6.61 for GCT, 26.68 for gender) are larger than the corresponding critical values $F_{0.99}^{\text{crit}} = 3.83$ and 7.31, respectively. The larger the F ratios, the more GCT and gender dif-

FIG. 8. Boxplots of the accuracies $\lambda_{i,s}$ averaged over all 10 mass elements ($i=1, \dots, 10$) at each plane ($s=1, \dots, 5$) for the 50 synthetic data sets. The mean values are marked with *. Higher performances occur at the superior planes, while lower values occur at the inferior planes.

fer. Also, the corresponding lower p-values ($3.5\text{E-}04 < 0.01$ for GCT, $7.0\text{E-}06 < 0.01$ for gender) provided sufficient evidence for such a conclusion. Additionally, for the analysis of κ and λ , no statistically significant differences based on GCT or gender, or the interaction between the two, were found, due to smaller F ratios ($< F_{0.95}^{\text{crit}}$) and larger p-values (> 0.05). Overall, analysis of variance for (Γ , κ , λ) confirmed the assumption (displayed in Table II) that GCT and gender partly influence (normalized error Γ) the optimization results.

B. Application to an *in vitro* hemilarynx experiment

In order to adapt the model dynamics to a hemilarynx data set, the initial rest positions of the mass elements were first estimated in accordance with the mean values of the 3D movements of the mounted marker points. Figure 9 shows the initial rest positions at one side of the 3DM. The average values of the computed optimization parameters ($\hat{Q}_p$, $\hat{Q}_{i,s}$, $\hat{Q}_r$) were 0.79 , 0.75 ± 0.36 , 0.94 ± 0.26 . The corresponding global measures (Γ , κ) were 0.16 and 84% , respectively. Accuracy of the optimization parameters could not be reported, since no predefined optimization parameters existed for the experimental data. However, the optimized subglottal pressure ($19.0 \text{ cm H}_2\text{O}$) and fundamental frequency (120 Hz) were sufficiently accurately reproduced.

Figure 10 shows the high similarity between the hemilarynx experimental 3D trajectories $\mathbf{c}_{3,3}[n]$ and adapted theoretical 3D trajectories $\mathbf{c}_{3,3}^M[n]$ of the mass element $m_{3,3}$, which responds to the marker-point located in the third vertical column and fourth sagittal line in the hemilarynx. The corresponding fundamental frequency was 120 Hz . The normalized errors Γ for trajectories in lateral, longitudinal, and vertical directions were 0.04 , 1.5 , and 0.26 , respectively. Their correlations κ were 99% , 68% , and 97% , respectively. The smallest error with highest correlation occurred in the lateral direction, while the highest error with lowest correlation occurred in the longitudinal direction.

Figure 11 shows the optimized model parameters applied to the hemilarynx experimental data set. The derived

FIG. 9. The rest positions of the 25 mass elements for one side are estimated in accordance to the mean values of the marker points placed on the hemilarynx. The marks $\times$ denote the rest position of the 25 mass elements.

FIG. 10. (Color online) Results of the adapted 3DM (dotted lines) of the mass element $m_{3,3}$ located in the median cross section at the plane $s=3$ on the right side of the model, compared to the hemilarynx trajectories (solid lines) at the corresponding suture-point. The fundamental frequency is 120 Hz.

masses $\hat{m}_{i,s}$ decreased from inferior to superior ($s=1$ to $s=5$) for all five coronal cross-sections [see Fig. 11(a)]. They ranged between 5.7×10^{-3} and 0.14 g. Additionally, from dorsal to ventral, the highest value occurred at the most ventral cross section $i=5$ for four planes ($s=2$ to $s=5$), except the most inferior plane $s=1$ (which had the highest value at $i=1$ near the vocal fold dorsal extreme). The lowest values at the superior planes ($s=4, s=5$) were at the most dorsal cross section $i=1$, while the inferior three planes ($s=1, 2, 3$) had the lowest values in cross section $i=4$. A large difference occurred at the transition from the most superior plane $s=5$ to the second plane $s=4$ [see Fig. 11(a)].

In Fig. 11(b), at the superior planes ($s=3$ to $s=5$), the anchor spring $\hat{k}_{i,s}^a$ decreased from inferior to superior for all five coronal cross-sections ($i=1$ to $i=5$). Moreover, the highest value occurred at the most inferior plane $s=1, i=4$. The lowest value was found at the most superior plane $s=5$ near the ventral extreme ($i=5$). Values of $\hat{k}_{i,s}^a$ at the two inferior planes ($s=1, 2$) were nearly equal at the dorsal cross-sections $i=1, 2$. Also, planes $s=2, 3$ had nearly equal values at $i=4$. The values ranged between 1.63 and 29.37 N/m.

Figure 11(c) shows that the longitudinal spring $\hat{k}_{i,s}^l$ had comparatively lower values at the most superior plane $s=5$.

For planes ($s=2, 5$), there was an increase from dorsal ($i=1, 2$) to ventral ($i=4, 5$). In general, the lowest value of $\hat{k}_{i,s}^l$ occurred at the most superior plane $s=5$ between the two dorsal cross-sections ($i=1, 2$), while the highest was found at the most inferior plane $s=1$ between cross-sections $i=2$ and $i=3$, shown in Fig. 11(c). The stiffnesses ranged between 17.98 and 1954.36 N/m.

In Fig. 11(d), the vertical spring $\hat{k}_{i,s}^v$ obviously decreased from inferior to superior for all five coronal cross-sections. A decrease occurred from dorsal to ventral for the superior spring ($s=4, 5$), and the inferior spring ($s=1, 2$). Values for the spring ($s=3, 4$) were nearly equal in all five coronal cross-sections. The highest value was found in the spring ($s=1, 2$) near the dorsal extreme ($i=1$), while the lowest was in $s=4, 5$ near the ventral extreme ($i=5$). The stiffnesses ranged between 24.53 and 1035.50 N/m.

IV. DISCUSSION

In this work, a mathematical optimization approach for approximating 3D vocal fold dynamics by a biomechanical model was suggested and verified. Validation of the optimization with fifty synthetic data sets and application to an *in*

FIG. 11. Logarithmic charts for the optimized model parameters after application to a hemilarynx experimental data set. (a)–(d) describe the mass $\hat{m}_{i,s}$, anchor stiffness $\hat{k}^a_{i,s}$, longitudinal stiffness $\hat{k}^l_{i,s}$, and the vertical stiffness $\hat{k}^v_{i,s}$ at/between different transverse planes s and coronal cross-sections i . $\hat{k}^v_{i,s}$ occurs between the current plane and the next upper plane. $\hat{k}^l_{i,s}$ occurs between the current plane and the next one. The achieved subglottal pressure $\hat{p}^{sub}$ was equal to 19.0 cm H₂O, and the estimated glottis length was 13.0 mm. $i = 1, \dots, 5$ denotes the coronal cross-sections from dorsal to ventral. $s = 1, \dots, 5$ denotes the transverse planes from inferior to superior. In general, the values of the model parameters at inferior plane ($s = 1$) are higher than those at superior plane ($s = 5$).

in vitro human hemilarynx experiment reflect its suitability and applicability. Among other things, the proposed method may be used to objectively quantify vocal fold asymmetries. A long-term goal of this method is to support clinical diagnosis and treatment of voice disorders through a better knowledge of the biomechanical parameters underlying the disorders.

A. Validation of optimization

1. Validation of feasibility via global measures

Using the proposed optimization procedure, the synthetic 3D trajectories were effectively approximated, as evidenced by the small normalized error ($\Gamma = 0.07 \pm 0.03$), the large correlation coefficient ($\kappa = 96 \pm 2\%$) and the high accuracy ($\lambda = 91 \pm 4\%$) over all cases, as shown in Table II. Since the optimization procedure is based on stochastic search-algorithms in which the calculated parameter values may not be unique, an ideal accuracy of 100% cannot be guaranteed for synthetic data. However, the derived measures (Γ , κ , λ) reflect the high reliability and robustness of the optimization procedure. Sufficiently high symmetry between right side and left side was reproduced. Also, it was shown that the performance of the optimization procedure was improved by using combinations of the PSO, SA, and PDSM algorithms. However, it was achieved at an expense of increased complexity of the optimization procedure.⁵⁴ Com-

pared to the verification procedure in Wurzbacher *et al.*,²⁸ the ranges of the measures (Sec. III A 1) in this work were reduced by more than 15%. Overall, the results confirm that the proposed optimization algorithms appropriately utilize dynamical information (e.g., from highspped imaging) to estimate biomechanical parameters of vocal folds.

2. Reproducibility of parameters

To further evaluate the inversion capability of the optimization, individual accuracies (λ_p , $\lambda_{i,s}$, λ_r) for different predefined optimization parameters (Q_p^S , $Q_{i,s}^S$, Q_r^S) after optimization were investigated (see Sec. III A). The differences among the individual accuracies (λ_p , $\lambda_{i,s}$, λ_r) may be explained by the fact that vocal fold dynamics may be more sensitive to changes in mass and stiffness than to changes in pressure and rest position. Additionally, relationships between the predefined parameters Q^S and the derived parameters $\hat{Q}$ are described in Fig. 7. High strengths of such relationships are demonstrated by the regression functions for $(Q_p, Q_{i,s}, Q_r)$, in Fig. 7, due to the high-percentage regression fits and the low disturbance terms. Also, these derived regression functions confirmed the high individual accuracies of $(Q_p, Q_{i,s}, Q_r)$ as shown in Sec. III A. Furthermore, compared to the whole value ranges of $(Q_p, Q_{i,s}, Q_r)$, the widths $\pm \Delta$ (see Fig. 7) of the 95% confidence intervals were relatively small. This shows that most of the variability in the relationships

between $\hat{Q}$ and Q^S occurs very close to the regression lines. Overall, these results reflect a high reproducibility of model parameters using the proposed optimization algorithms.

3. Influence from weighting coefficients

Since the observed vocal fold 3D dynamics along the glottal surface are anisotropic, a concept of the weighting coefficients (g_μ^d, g_s^p, g_i^m , see Sec. II B 2) was utilized in the optimization procedure. The estimated weighting coefficients [Eqs. (16) and (17)] reflect the relative amplitudes of vibration on different parts of the vocal fold. In Fig. 8 it was shown that the individual accuracies $\lambda_{i,s}$ were different between planes. In particular, the variation in the accuracies corresponded to the variations in weighting coefficient g_s^p for the different planes [see Eq. (15)]. That is, higher values of weighting coefficients ($g_3^p = g_4^p = 0.3$) yielded comparatively better individual accuracies for the corresponding planes. A reason might be that the parameter optimization procedure was influenced by the weighting coefficients g_s^p . In other words, it can be inferred that the weighting coefficients (g_μ^d, g_s^p, g_i^m) may be used to help guide the optimization results. Or put another way, by means of the weighting coefficients, the optimization procedure can be designed to selectively focus on portions of the vocal folds deemed most significant for a particular application, such as where the greatest amplitudes of vibration occur, or where diseased or scarred portions of vocal fold tissue may occur.

4. GCT and gender

In determining the effects of GCT and gender on the optimization results, one significant issue is how the interaction of these two variables may influence the results. Analysis of variance revealed that both factors significantly affected the objective function Γ , while there was not sufficient evidence to suggest that the correlation coefficient κ or the accuracy λ was significantly influenced by these factors (see Sec. III A 4). This conclusion might be explained by the fact that the initial setup of GCT may implicitly influence the normalized error between adapted and synthetic dynamics, since mass distributions of the model differ for GCTs. Moreover, synthetic female data sets, which exhibit higher fundamental frequencies and larger sub-glottal pressures,⁴⁵ yield larger Γ s (see Table II). This may be attributed to the fact that larger subglottal pressures may result in short-term irregularities in the 3D dynamics of the model, since higher subglottal pressures are more likely to induce irregular and asymmetric vibrations.^{60,61} Therefore, for larger subglottal pressures, the distance between synthetic and adapted 3D trajectories may be increased. However, the results of the analysis of variance for (Γ, κ, λ) show that neither factor significantly affected either dynamic correlation or optimization accuracies.

B. Application to an *in vitro* hemilarynx experiment

Since human vocal fold 3D dynamics are governed by biomechanical tissue properties which can be expressed in

terms of quantitative parameters, it is important to examine whether the physiologically relevant properties can be accurately extracted using the 3DM with the proposed optimization procedure. For this purpose, the optimization procedure was also applied to an experimental hemilarynx data set. Compared to optimization results for synthetic data sets, the correlation coefficient ($\kappa = 84\%$) was lower, while the corresponding normalized error ($\Gamma = 0.16$) was higher. This might be attributed to two causes. One reason is that the boundary conditions, such as the 3D positions of the ventral/dorsal vocal fold extremes are not known exactly, and the corresponding estimates of such fixed positions might not be sufficiently close to the actual values. Another reason is that the recorded experimental curves might not be sufficiently smooth as shown in Fig. 10, due to the image processing errors (up to 20%). However, the optimization results reasonably fulfilled the combined condition Eq. (20). It exhibits a successful optimization, and reveals a sufficient similarity between the adapted model dynamics and the hemilarynx dynamics.

As an example of an optimization result, the adapted 3D trajectories of the mass element $m_{3,3}$ and corresponding hemilarynx experimental 3D trajectories were compared in Fig. 10. This supported the evidence that values of normalized errors Γ and correlations κ were steered by the weighting coefficients g_μ^d for different displacement directions. Relatively lower normalized errors with higher correlations were derived in the lateral and vertical directions, since the corresponding weighting coefficients were explicitly higher [$g_x^d = 0.5, g_z^d = 0.4$ in Eq. (16)]. The values of Γ and κ in the longitudinal direction were less accurate, due to the lower value of the weighting coefficient for the longitudinal direction ($g_y^d = 0.1$). However, longitudinal displacements are usually considered a relatively minor aspect of 3D vocal fold dynamics. In other words, at the dimension where a higher value of the weighting coefficient appears, a lower Γ with higher κ can be achieved.

Overall, the derived stiffness ($\hat{k}_{i,s}^a, \hat{k}_{i,s}^l, \hat{k}_{i,s}^v$) increased from superior to inferior [see Figs. 11(b)–11(d)]. These results support the suggestion proposed by Story *et al.*⁶² that vocal fold vibration can best be modeled if the stiffness of the superior parts of the vocal fold is smaller than that of the inferior parts. In addition, from an anatomic point of view, the decrease [Fig. 11(a)] of the mass $\hat{m}_{i,s}$ from inferior to superior could be explained by the fact that the inferior parts of the vocal fold are closer to the stiff tissue of the trachea as well as the cricoid cartilage. The relatively high differences in the stiffness $\hat{k}_{i,s}^a, \hat{k}_{i,s}^l$, and $\hat{k}_{i,s}^v$ from inferior to superior implicitly reflect that relatively significant variations in tissue environments occur along the vertical direction. Also, either the ventral or dorsal cross-sections had slightly higher stiffness/mass values. This could be due to the locations of the ventral/dorsal cross-sections which are close to the vocal fold fixed endings. Similar to the conclusions in Berry *et al.*⁶³ and Alipour *et al.*,⁶⁴ a finding was that the longitudinal stiffness $\hat{k}_{i,s}^l$ was stronger than the stiffness in the other two directions (i.e., lateral and vertical). This nonisotropic nature of the stiffness might be regarded as an important consideration for phonosurgeons who wish to avoid

subluxation of the cricoarytenoid joint in patients, as suggested in Ref. 63.

C. Performance of optimization procedure and outlook

Extending the optimization procedure of the 2DM,³⁰ the 3DM and resulting optimization parameters are computed in a 3D space to more realistically investigate the physiological properties of the vocal folds.²⁴ Based on the hemilarynx experimental data and laryngeal physiological structure,⁴ 50 mass-spring oscillator units are coupled within the 3DM. With this greater complexity, computational costs of the optimization procedure are significantly increased. However, it is significantly more useful for analyzing the 3D dynamics of vocal folds. For parameter optimization of a 175 millisecond segment of sustained synthetic 3D dynamics, the required optimization time is currently approximately 3–5 days on a standard PC [Intel(R) Core(TM)2 Duo E8500, 3.16 GHz, 3.50 GB RAM, C#]. The optimization procedure was also successfully applied to a hemilarynx experimental data set. These results provide an important first step in fitting 3D vocal fold dynamics, in preparation for larger systematic studies of both *in vitro* and *in vivo* experiments. In the future, the proposed 3DM optimization approach may be further adapted to investigate specific voice disorders, studying specific pathological regions of the vocal folds, perhaps with especially designed weighting coefficients to focus on the region of interest.

On the whole, this study shows that the 3DM with its optimization procedure enables modeling of human vocal fold 3D dynamics. The model parameters capture the biomechanical properties of vocal fold vibrations. Thus, the adapted model parameters may be used to further disclose essential characteristics of vocal fold dynamics, not only in planar forms, but also in spatial forms. The future goal is the clinical application: to adapt the model toward vocal fold 3D dynamics, which are extracted from *in vivo* human laryngeal recordings using endoscopic HS imaging in combination with a laser-spot projection system (necessary for computation of 3D tissue coordinates). The left-right asymmetry quantification for local vocal fold parts (five positions along the vocal fold length) may play a vital role in objective judgments of surgical interventions or conventional laryngeal therapy. Additionally, from a clinical perspective, the consequences of surgical treatment (i.e., scarring) resulting in different spring and damping components can be interpreted and analyzed by future parameter studies of the 3DM.

V. SUMMARY

The aim of the proposed 3DM parameter optimization procedure was to provide an opportunity to objectively quantify biomechanical properties of 3D human vocal fold vibrations. Investigating the relationships between model parameters and vocal fold dynamics may be beneficial to voice therapy and surgical treatments of voice disorders. Due to non-convexity of the optimization problem, different global and local optimization algorithms were applied to yield appropriate results. To validate the reliability and feasi-

bility of the optimization procedure, different synthetic data sets were adapted with five different glottal closure types and genders. The results of the optimization procedure were further investigated using data from an *in vitro* hemilarynx experiment. In future studies, both *in vitro* and *in vivo* hemilarynx data will be systematically adapted to conduct an in-depth investigation regarding the correspondence between biomechanical tissue parameters and the resulting vocal fold vibrations.

Finally, the proposed optimization strategy need not be confined to the study of vocal fold vibration, but is sufficiently general that it may also be applied to other fields of interest.

ACKNOWLEDGMENTS

This work was supported by DFG (*Deutsche Forschungsgemeinschaft*, German Research Foundation) Grant No. FOR 894/1 “Strömungsphysikalische Grundlagen der Menschlichen Stimmgebung” (i.e., fluid dynamical basics of human voice production). Support for D.A.B. on this project was funded by NIH/NIDCD Grant No. R01 DC03072.

¹M. Hirano, “Phonosurgery: Basic and clinical investigations,” *Otologia* (Fukuoka) **21**(Suppl. 1), 239–262 (1975).

²T. Baer, “Investigation of phonation using excised larynxes,” in Ph.D. dissertation (Massachusetts Institute of Technology, Cambridge, MA, 1975).

³M. Döllinger, D. A. Berry, and G. S. Berke, “A quantitative study of the medial surface dynamics of an *in vivo* canine vocal fold during phonation,” *Laryngoscope* **115**(9), 1646–1654 (2005).

⁴A. Boessenecker, D. A. Berry, J. Lohscheller, U. Eysholdt, and M. Döllinger, “Mucosal wave properties of a human vocal fold,” *Acta Acust. Acust.* **93**(9), 815–823 (2007).

⁵A. Sonninen and A.-M. Laukkanen, “Hypothesis of whiplike motion as a possible traumatizing mechanism in vocal fold vibration,” *Folia Phoniatr. Logop.* **55**, 189–198 (2003).

⁶R. Baken, “Irregularity of vocal period and amplitude: A first approach to the fractal analysis of voice,” *J. Voice* **4**, 185–197 (1990).

⁷M. Döllinger, U. Hoppe, F. Hettlich, J. Lohscheller, S. Schubert, and U. Eysholdt, “Vibration parameter extraction from endoscopic image series of the vocal folds,” *IEEE Trans. Biomed. Eng.* **49**(8), 773–781 (2002).

⁸N. Isshiki, M. Tanabe, K. Ishizaka, and D. Broad, “Clinical significance of asymmetrical vocal cord tension,” *Ann. Otol. Rhinol. Laryngol.* **86**(1), 58–66 (1977).

⁹S. Niimi and M. Miyaji, “Vocal fold vibration and voice quality,” *Folia Phoniatr. Logop.* **52**(1–3), 32–38 (2000).

¹⁰J. 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**(6), 3179–3192 (2001).

¹¹I. R. Titze, *Principles of Voice Production* (Prentice Hall, Englewood Cliffs, NJ, 1994), pp. 23–52, 307–322.

¹²I. Titze, R. Baken, and H. Herzel, in *Vocal Fold Physiology: Frontiers in Basic Science*, edited by I. Titze (Singular Publishing Group, San Diego, CA, 1993), pp. 143–188.

¹³A. E. Aronson and D. M. Bless, in *Clinical Voice Disorders*, edited by T. Y. Hiscock (Thieme Medical Publisher, New York, 2009), pp. 2–9.

¹⁴Q. Fujimura, in *Vocal Fold Physiology*, edited by K. N. Stevens and M. Hirano (University of Tokyo Press, Tokyo, 1981), pp. 271–288.

¹⁵I. R. Titze, in *The Myoelastic Aerodynamic Theory of Phonation*, edited by S. Klemuk (National Center for Voice and Speech, Iowa City, IA, 2006), pp. 1–62.

¹⁶S. Hadjitodorov and P. Mitav, “A computer system for acoustic analysis of pathological voices and laryngeal diseases screening,” *Med. Eng. Phys.* **24**(6), 419–429 (2002).

¹⁷S. Fleischer and M. Hess, “The significance of videos in laryngological practice,” *HNO* **54**, 628–634 (2006).

¹⁸J. Wendler, “Stroboscopy,” *J. Voice* **6**(2), 149–154 (1992).

- ¹⁹J. G. Švec and H. K. Schutte, "Videokymography: High-speed line scanning of vocal fold vibration," *J. Voice* **10**(2), 201–205 (1996).
- ²⁰R. Schwarz, M. Döllinger, T. Wurzbacher, U. Eysholdt, and J. Lohscheller, "Spatio-temporal quantification of vocal fold vibrations using high-speed videendoscopy and a biomechanical model," *J. Acoust. Soc. Am.* **123**, 2717–2732 (2008).
- ²¹M. Döllinger, "The next step in voice assessment: High-speed digital endoscopy and objective evaluation," *Curr. Bioinf.* **4**(2), 101–111 (2009).
- ²²T. Braunschweig, P. Schelhorn-Neise, and M. Döllinger, "Diagnosis of functional voice disorders by using the high speed recording technics," *Laryngorhinotologie* **87**(5), 323–330 (2008).
- ²³D. D. Deliyski, P. P. Petrushev, H. S. Bonilha, T. T. Gerlach, B. Martin-Harris, and R. E. Hillman, "Clinical implementation of laryngeal high speed videendoscopy: Challenges and evolution," *Folia Phoniatr. Logop.* **60**(1), 33–44 (2008).
- ²⁴M. Döllinger, D. A. Berry, and D. W. Montequin, "The influence of epilarynx area on vocal fold dynamics," *Otolaryngol. Head. Neck. Surg.* **135**(5), 724–729 (2006).
- ²⁵M. Döllinger and D. A. Berry, "Visualization and quantification of the medial surface dynamics of an excised human vocal fold during phonation," *J. Voice* **20**(3), 401–413 (2006).
- ²⁶A. Yang, D. A. Berry, J. Lohscheller, D. Voigt, U. Eysholdt, and M. Döllinger, "Biomechanical modeling of the three-dimensional aspects of human vocal fold dynamics," *J. Acoust. Soc. Am.* **127**(2), 1014–1031 (2010).
- ²⁷X. Zheng, S. Bielamowicz, H. Luo, and R. Mittal, "A computational study of the effect of false vocal folds on glottal flow and vocal fold vibration during phonation," *Ann. Biomed. Eng.* **37**(3), 625–641 (2009).
- ²⁸T. Wurzbacher, R. Schwarz, M. Döllinger, U. Hoppe, U. Eysholdt, and J. Lohscheller, "Model-based classification of nonstationary vocal fold vibrations," *J. Acoust. Soc. Am.* **120**(2), 1012–1027 (2006).
- ²⁹R. Schwarz, U. Hoppe, M. Schuster, T. Wurzbacher, U. Eysholdt, and J. Lohscheller, "Classification of unilateral vocal fold paralysis by endoscopic digital high-speed recordings and inversion of a biomechanical model," *IEEE Trans. Biomed. Eng.* **53**(6), 1099–1108 (2006).
- ³⁰T. Wurzbacher, M. Döllinger, R. Schwarz, U. Hoppe, U. Eysholdt, and J. Lohscheller, "Spatio-temporal classification of vocal fold dynamics by a multi-mass model comprising time-dependent parameters," *J. Acoust. Soc. Am.* **123**, 2324–2334 (2008).
- ³¹M. Döllinger and D. A. Berry, "Computation of the three-dimensional medial surface dynamics of the vocal folds," *J. Biomech.* **39**(2), 369–374 (2006).
- ³²M. Döllinger, N. Tayama, and D. A. Berry, "Empirical eigenfunctions and medial surface dynamics of a human vocal fold," *Methods Inf. Med.* **44**(3), 384–391 (2005).
- ³³G. Luegmair, S. Kniesburges, M. Zimmermann, Alexander, Sutor, U. Eysholdt, and M. Döllinger, "Optical reconstruction of high-speed surface dynamics in an uncontrollable environment," *IEEE Trans. Med. Imaging* **29**(12), 1979–1991 (2010).
- ³⁴J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of the IEEE International Conference on Neural Networks*, Perth, WA, Australia (IEEE Service Center, Piscataway, NJ, 1995), Vol. 4, pp. 1942–1948.
- ³⁵L. Ingber, "Adaptive simulated annealing," *J. Control Cybernetics* **25**(1), 33–54 (1996).
- ³⁶W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, New York, 1994), pp. 348–455.
- ³⁷J. Lohscheller, U. Eysholdt, H. Toy, and M. Döllinger, "Phonovibrography: Mapping high-speed movies of vocal fold vibrations into 2-d diagrams for visualizing and analyzing the underlying laryngeal dynamics," *IEEE Trans. Med. Imaging* **27**, 300–309 (2008).
- ³⁸O. Rasp, J. Lohscheller, M. Döllinger, U. Eysholdt, and U. Hoppe, "The pitch rise paradigm: A new task for real-time endoscopy of non-stationary phonation," *Folia Phoniatr. Logop.* **58**, 175–185 (2006).
- ³⁹P. Dejonckere, P. Bradley, P. Clemente, G. Cornut, L. Crevier-Buchman, G. Friedrich, P. V. D. Heyning, M. Remacle, V. Woisard, and Committee on Phoniatrics of the European Laryngological Society (ELS), "A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. guideline elaborated by ELS," *Eur. Arch. Otorhinolaryngol.* **258**(2), 77–82 (2001).
- ⁴⁰T. Ikeda, Y. Matsuzaki, and T. Aomatsu, "A numerical analysis of phonation using a two-dimensional flexible channel model of the vocal folds," *J. Biomech. Eng.* **123**(6), 571–579 (2001).
- ⁴¹K. Ishizaka and J. L. Flanagan, "Synthesis of voiced sounds from a two-mass model of the vocal cords," *Bell Syst. Techn. J.* **51**(6), 1233–1268 (1972).
- ⁴²R. W. Chan, "Vocal fold tissue failure: Preliminary data and constitutive modeling," *J. Biomech. Eng.* **126**(4), 466–474 (2004).
- ⁴³I. R. Titze, "The physics of small-amplitude oscillation of the vocal folds," *J. Acoust. Soc. Am.* **83**(4), 1536–1552 (1988).
- ⁴⁴I. Titze and P. Durham, in *Laryngeal Function in Phonation and Respiration*, edited by T. Baer, C. Sasaki, and K. Harris (College Hill, San Diego, CA, 1986).
- ⁴⁵L. P. Fulcher, R. C. Scherer, A. Melnykov, V. Gateva, and M. E. Limes, "Negative coulomb damping, limit cycles, and self-oscillation of the vocal folds," *Am. J. Phys.* **74**(5), 386–393 (2006).
- ⁴⁶E. Cataldo, C. Soize, R. Sampaio, and C. Desceliers, "Probabilistic modeling of a nonlinear dynamical system used for producing voice," *Comput. Mech.* **43**, 265–275 (2009).
- ⁴⁷D. A. Berry, D. W. Montequin, and N. Tayama, "High-speed digital imaging of the medial surface of the vocal folds," *J. Acoust. Soc. Am.* **110**(5), 2539–2547 (2001).
- ⁴⁸J. Kobayashi, E. Yumoto, M. Hyodo, and K. Gyo, "Two-dimensional analysis of vocal fold vibration in unilaterally atrophied larynges," *Laryngoscope* **110**, 440–446 (2000).
- ⁴⁹M. Döllinger, D. A. Berry, and G. S. Berke, "Medial surface dynamics of an in vivo canine vocal fold during phonation," *J. Acoust. Soc. Am.* **117**(5), 3174–3183 (2005).
- ⁵⁰M. Spencer, T. Siegmund, and L. Mongeau, "Determination of superior surface strains and stresses, and vocal fold contact pressure in a synthetic larynx model using digital image correlation," *J. Acoust. Soc. Am.* **123**(2), 1089–1103 (2008).
- ⁵¹J. F. Schutte, B.-I. Koh, J. A. Reinbolt, R. T. Haftka, A. D. George, and B. J. Fregly, "Evaluation of a particle swarm algorithm for biomechanical optimization," *J. Biomech. Eng.* **127**(3), 465–474 (2005).
- ⁵²M. Khanesar, M. Teshnehlav, and M. Shooreshdeli, "A novel binary particle swarm optimization," in *Proceedings of the 15th Mediterranean Conference on Control and Automation* [Mediterranean Control Association (MED)], Athens, Greece, 2007, Vol. T33–001, pp. 1–6.
- ⁵³L. Ingber and B. Rosen, "Genetic algorithms and very fast simulated reannealing: A comparison," *Mathl. Comput. Modelling* **16**(11), 87–100 (1992).
- ⁵⁴*Particle Swarm Optimization*, edited by A. Lazinic (In-Tech Education and Publishing, Vienna, Austria, 2009), pp. 155–168.
- ⁵⁵M. Schuster, J. Lohscheller, P. Kummer, U. Eysholdt, and U. Hoppe, "Laser projection in high-speed glottography for high-precision measurements of laryngeal dimensions and dynamics," *Eur. Arch. Otorhinolaryngol.* **262**(6), 477–481 (2005).
- ⁵⁶M. P. Gelfer and D. K. Bultemeyer, "Evaluation of vocal fold vibratory patterns in normal voices," *J. Voice* **4**, 335–345 (1990).
- ⁵⁷I. R. Titze, "Physiologic and acoustic differences between male and female voices," *J. Acoust. Soc. Am.* **85**(4), 1699–1707 (1989).
- ⁵⁸A. M. Sulter and F. W. J. Albers, "The effects of frequency and intensity level on glottal closure in normal subjects," *Clin. Otolaryngol. Allied Sci.* **21**, 324–327 (1996).
- ⁵⁹A. M. Sulter, H. K. Schutte, and D. G. Miller, "Standardized laryngeal videostroboscopic rating: Differences between untrained and trained male and female subjects, and effects of varying sound intensity, fundamental frequency and age," *J. Voice* **10**(2), 175–189 (1996).
- ⁶⁰J. J. Jiang, Y. Zhang, and J. Stern, "Modeling of chaotic vibrations in symmetric vocal folds," *J. Acoust. Soc. Am.* **110**(4), 2120–2128 (2001).
- ⁶¹Y. Zhang and J. J. Jiang, "Asymmetric spatiotemporal chaos induced by a polypoid mass in the excised larynx," *Chaos* **18**(4), 043102 (2008).
- ⁶²B. H. Story and I. R. Titze, "Voice simulation with a body-cover model of the vocal folds," *J. Acoust. Soc. Am.* **97**(2), 1249–1260 (1995).
- ⁶³D. A. Berry, D. W. Montequin, R. W. Chan, I. R. Titze, and H. T. Hoffman, "An investigation of cricoarytenoid joint mechanics using simulated muscle forces," *J. Voice* **17**, 47–62 (2003).
- ⁶⁴F. Alipour-Haghighi, D. A. Berry, and I. R. Titze, "A finite-element model of vocal fold vibration," *J. Acoust. Soc. Am.* **108**(6), 3003–3012 (2000).