# Biomechanical modeling of the three-dimensional aspects of human vocal fold dynamics

Anxiong Yang[a)]
*Department of Phoniatrics and Pediatric Audiology, University Hospital Erlangen, Medical School,*
*Bohlenplatz 21, 91054 Erlangen, Germany*

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

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

Stefan Becker
*Institute of Fluid Mechanics, Friedrich-Alexander University of Erlangen-Nuremberg, Cauerstrasse 4,*
*91058 Erlangen, Germany*

Ulrich Eysholdt, Daniel Voigt, and Michael Döllinger
*Department of Phoniatrics and Pediatric Audiology, University Hospital Erlangen, Medical School,*
*Bohlenplatz 21, 91054 Erlangen, Germany*

Human voice originates from the three-dimensional (3D) oscillations of the vocal folds. In previous studies, biomechanical properties of vocal fold tissues have been predicted by optimizing the parameters of simple two-mass-models to fit its dynamics to the high-speed imaging data from the clinic. However, only lateral and longitudinal displacements of the vocal folds were considered. To extend previous studies, a 3D mass-spring, cover-model is developed, which predicts the 3D vibrations of the entire medial surface of the vocal fold. The model consists of five mass planes arranged in vertical direction. Each plane contains five longitudinal, mass-spring, coupled oscillators. Feasibility of the model is assessed using a large body of dynamical data previously obtained from excised human larynx experiments, *in vivo* canine larynx experiments, physical models, and numerical models. Typical model output was found to be similar to existing findings. The resulting model enables visualization of the 3D dynamics of the human vocal folds during phonation for both symmetric and asymmetric vibrations.
© 2010 Acoustical Society of America. [DOI: 10.1121/1.3277165]

## I. INTRODUCTION

Human voice is generated by the oscillations of two lateral opposing vocal folds located in the larynx.[1] The vocal fold oscillations are excited by air flow through the trachea generated by the lung. Moreover, the characteristics of the acoustic signal, e.g., sound pressure level and pitch of tone, are influenced by vocal fold tensions.[2] A disturbed or hoarse voice arises from irregular vocal fold vibrations.[3] These irregularities are often caused by anatomical asymmetries between left and right vocal fold.[3–5] Asymmetries in the larynx such as unilateral vocal fold polyps or paralysis can be directly observed using endoscopes.[1] However, vocal fold pathology such as functional dysphonia, which presents no visual evidences for morphological laryngeal abnormalities, is only visible during vocal fold oscillation.[1] Hence, analyzing vocal dynamics is a critical component of voice research.

In the clinic, endoscopic high-speed (HS) digital imaging of the vocal fold vibration (100–300 Hz) is becoming increasingly more common, typically with sampling rates of 2000–4000 fps (frames per second). An appropriate approach to analyze endoscopically recorded vocal fold dynamics is the adaption of biomechanical models.[1] Therefore, one or more trajectories at specific locations on the vocal fold edges (see Fig. 1 in Ref. 6) are reproduced. Due to the increase in computational power, this method has been gradually improved:

First, only one lateral trajectory was recreated by using a two-mass-model (2MM).[7,8] The corresponding optimization was manually performed.[9] The first fully automatic optimization procedure was realized by Döllinger *et al.*[1] applying the Nelder–Mead algorithm. With aids of this parameter optimization procedure, the dynamics of healthy voices were simulated.[10] To differentiate normal from unilateral vocal

---

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

fold paralysis, classification schemes were proposed.[2] Moreover, according to clinical diagnosis some specific laryngeal diseases exhibit their clinical picture only at certain vocal fold oscillation frequencies; therefore, parameter optimization for non-stationary vocal fold dynamics was performed.[11]

Later, vocal fold movements at three positions (see Fig. 1 in Ref. 6) located along the longitudinal direction were adapted by enhanced two-dimensional (2D) multi-mass-models[6,12] (2DMs). The optimization procedure for sustained phonation was successfully accomplished through verification on two normal and six pathological voices.[6] In addition, for the corresponding non-stationary phonation, a so-called block-based optimization procedure was developed and verified with clinical data sets.[12]

To date, most low-dimensional, biomechanical models of the vocal folds have considered vibration in only one or two dimensions.[13–18] Similarly, optimization has focused primarily on the analysis of superior dynamics of the vocal folds,[6] from a superior aspect. However, it is important to objectively quantify vertical movements[19] as well as medial surface dynamics of the vocal folds.[20,21] Moreover, mucosal wave propagation, which occurs primarily in vertical direction, establishes a continually mechanical transfer of energy from the glottal airstream to the oscillating tissue of vocal folds.[22–24] It is noteworthy that such models could be unduly restricted, if the most complicated aspects of vocal fold dynamics, which occur along the medial surface (including coupling with the glottal air, collision with the opposing vocal fold, etc.),[20,25] are not reasonably taken into account. Therefore, fully analyzing the vertical vibrations,[21] the medial surface dynamics of the folds, as well as the mucosal wave propagation,[19,26] is essential to quantify the vocal fold mechanical characteristics.[20]

In this work, we develop a 3D mass-spring, cover-model of vocal fold vibration. This model predicts the 3D vibrations of the entire medial surface of the vocal fold. In this work, we will assess the feasibility of the model using a large amount of dynamical data obtained from laboratory larynx experiments. Ultimately, one of the primary purposes of this model is to infer an appropriate approximation of biomechanical tissue properties across the entire medial surface of the vocal folds, as the model parameters are optimized to fit model dynamics to the experimental data of vocal fold vibrations. Therefore, it is hoped that the resultant vocal fold dynamics may be firmly interpreted in a physiological way. The model was developed to meet the following criteria.

## A. 3D dynamics and mucosal wave

In order to simulate 3D vocal fold dynamics as realistic as possible, the findings of several 3D model templates were taken into account: human hemilarynx,[19,21,26–28] *in vivo* canine dynamics,[29,30] excised canine larynges,[31–34] physical models,[35] synthetic models[36,37] as well as numerical finite-element models.[38,39]

Experiments with excised canine larynges reflect that the amplitudes of the elliptical paths traversed by flesh-points arranged from superior to inferior are decreased.[32] The largest displacements and velocities were concentrated in the up-
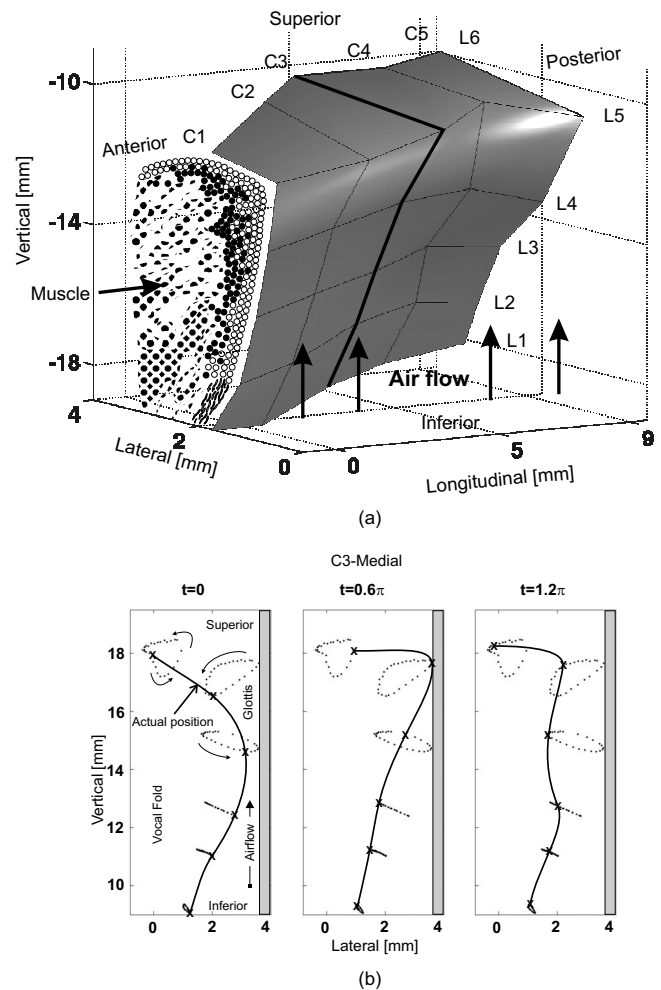


(a)



(b)

FIG. 1. (a) 3D geometry of one vocal fold surface obtained from hemilarynx experiments (Ref. 26). 30 surgical sutures serving as tracking markers are homogeneously distributed on the vocal fold surface. (Ref. 26). (b) Movements of the tracking markers along the vertical column C3 during one oscillation cycle. The antidromic vibration in lateral direction between inferior and superior sutures is clearly displayed (Ref. 26).

per medial portion of *in vivo* canine vocal folds.[30]

The human hemilarynx set-up was also considered. It allows to observe the dynamics of an entire vocal fold surface, since the other vocal fold is cut off.[21,27–29] For visualizing 3D trajectories extracted from performed HS recordings, six rows from inferior to superior and five vertical columns from ventral to dorsal of surgical micro-sutures were uniformly distributed over the vocal fold surface,[26] see Fig. 1(a). A typical oscillation cycle along the vertical suture column (C3) is shown in Fig. 1(b). The cycle started when the most inferior suture was at the most medial position and begins to move in lateral direction.[26] Within these studies, local 3D displacements, velocities, and accelerations were described and analyzed.[19,26]

During phonation the dynamics of human vocal fold mucosa occur and propagate wavelike in anterior-posterior, medio-lateral, and in caudal-cranial direction.[40] Bending and twisting may take place between the faster and the more slowly moving portions of the vocal folds in the lateral, longitudinal, and vertical dimensions, respectively.[40]

**Entire model**



**Vertical cross-section**          **Horizontal plane**
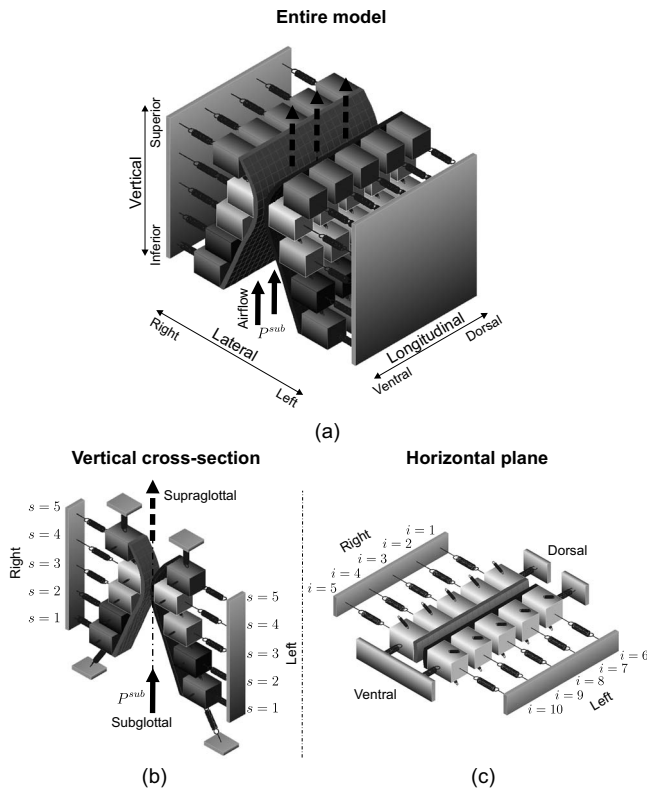


(b)          (c)

FIG. 2. (a) Schematic three-dimensional representation of the 3DM. Each of both vocal folds consists of five masses on each horizontal plane. The masses are elastically connected to a rigid body in lateral direction by using anchor springs. In addition, the masses are connected to each other through springs located in vertical and longitudinal directions. (b) 3D view of the individual cross-section of the 3DM with plane index $s$. (c) 3D view of one individual plane with column index $i$. The rigid bodies representing the anterior commissure and the fixed positions at the dorsal end of both vocal folds are displayed.

## B. Different glottis closure types

Endoscopic HS digital recordings[1,6,12] were taken as a basis to simulate the five common glottis closure types during normal phonation. The different glottis closure types have been described in detail before[41–44] and are, respectively, classified as *rectangle, hourglass, triangular-pointed dorsal, triangular-pointed ventral, and convex* (see Fig. 5 in Ref. 41).

## II. METHODS

### A. 3D-multi-mass-model

The vibrational function of human vocal folds depend on myoelastic and aerodynamic properties of the larynx,[1] which are simulated rationally as far as possible within the proposed 3D-multi-mass-model (3DM). The 3DM is on the basis of previous 2DM.[45,46] A corresponding three-dimensional view is sketched in Fig. 2.

Each vocal fold is supposed to consist of five horizontal layers (planes) arranged from inferior to superior. Corresponding to the hemilarynx experiments, each plane contains five longitudinal coupled mass-spring oscillators, see Fig. 2. The masses situated at different vertical planes are elastically connected by using springs in vertical direction. At the ven-
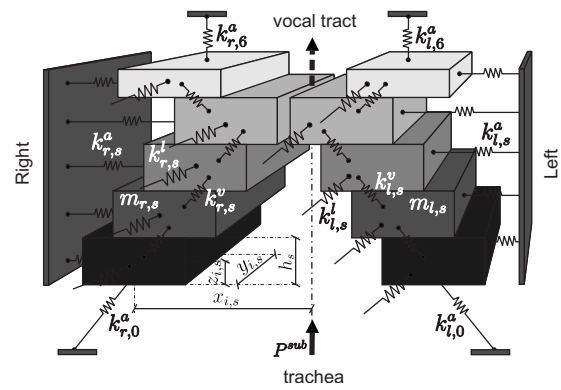


FIG. 3. Schematic representation of the cross-section of the 3DM of the vocal folds. The stiffness of anchor springs, longitudinal springs and vertical springs are indicated with $k_{r,s}^a$, $k_{r,s}^l$, $k_{r,s}^v$ with $s=1,\ldots,5$ from inferior to superior. The subscript $r$ means the lateral right side of the vocal fold ($i=1,\ldots,5$). Likewise, the stiffness of different springs on the lateral left side are symbolized as $k_{l,s}^a$, $k_{l,s}^l$, $k_{l,s}^v$, with index $i=6,\ldots,10$. The thickness of the mass element located in plane $s$ is labeled as $h_s$. The position of the mass element $m_{i,s}$ is denoted with 3D-components $x_{i,s}$, $y_{i,s}$, $z_{i,s}$. The damping $r_{i,s}$ for each mass element is not shown so as to keep clarity of the drawing.

tral end of the model, the left lateral and right lateral sides of the 3DM are connected together at an anchor being similar to the anterior commissure of the larynx, where both vocal folds merge together,[46] Fig. 2(c). Additionally, the dorsal extremities of both sides of the 3DM are coupled to their own fixed positions that are joined with the arytenoid cartilages, which do not vibrate during phonation.[46]

The viscous losses inside the glottis during vocal fold dynamics are neglected.[7,8,47] During phonation the driving aerodynamic force acting on left and right vocal fold is caused by the subglottal pressure $P^{\text{sub}}$ which follows the flow dynamic Bernoulli equation.[1,48,49] Thus, under the action of subglottal pressure, the masses located on both sides are set into vibration. Because of simplicity of the 3DM, the nonlinearities[7] of the elastic forces are small enough to be negligible.[1]

The tissue properties of the vocal folds are simplified and represented with lumped elements, namely, masses $m_{i,s}$, dampings $d_{i,s}$, and different associated spring stiffness coefficients, e.g., anchor spring stiffness $k_{i,s}^a$, vertical stiffness $k_{i,s}^v$, and longitudinal stiffness $k_{i,s}^l$ (Fig. 3). The indices $(i,s)$ label the different mass element ($i=1,\ldots,5$ on the right side, and $i=6,\ldots,10$ on the left side, from dorsal to ventral) at different planes with $s=1,\ldots,5$ from inferior to superior, see Fig. 2. The length of the model in the longitudinal direction [from the fixed position at the ventral end to the fixed position at the dorsal end, see Fig. 2(c)] is defined as the glottal length $l_g$ (approximately 11–15 mm) of human vocal folds, based on the Su *et al.*[50] and Gray's anatomy.[51] The height of the model from the most inferior mass element to the most superior mass element is 10 mm. In the lateral direction the distance between the mass element and the glottal midline is 0.425 mm for minimum, and 2.425 mm for maximum. The more specific physical dimensions (e.g., vertical and longitudinal spacings between the mass elements) are referenced in Table I in Sec. II B.

TABLE I. Brief summary of initial model parameters of the 3DM for the simulation of 3D vocal fold dynamics. $i=1,\ldots,10$ is the index of each mass element at each plane. $s=1,\ldots,5$ denotes the plane from inferior to superior. $\widetilde{k}_{i,s}^v$ occurs between the current plane $s$ and the next upper plane $s+1$. Moreover, $\widetilde{k}_{i,6}^v$, $\widetilde{k}_{i,0}^a$ indicating the stiffness of the upper and lower vertical anchor springs illustrated in Fig. 3 are defined with the values 5 and 450 N/m respectively.

| Plane $s$ | $\widetilde{m}_{i,s}$ (g) | $\widetilde{k}_{i,s}^a$ (N/m) | $\widetilde{k}_{i,s}^v$ (N/m) | $h_s$ (mm) | $\|\widetilde{x}_{i,s}^r\|$ (mm) | $\widetilde{z}_{i,s}^r$ (mm) |
|---|---|---|---|---|---|---|
| 1 | 0.02 | 50 | 450 | 2 | 1.675 | 0 |
| 2 | 0.02 | 20 | 80 | 2 | 1.075 | $0.5(h_1+h_2)$ |
| 3 | 0.02 | 20 | 50 | 2 | 0.475 | $\widetilde{z}_{i,2}^r+0.5(h_2+h_3)$ |
| 4 | 0.02 | 8 | 50 | 2 | 0.425 | $\widetilde{z}_{i,3}^r+0.5(h_3+h_4)$ |
| 5 | 0.01 | 1.6 | $\cdots$ | 1 | 2.425 | $\widetilde{z}_{i,4}^r+0.5(h_4+h_5)$ |
| Cross-section | $i=5,10$ | $i=4,9$ | $i=3,8$ | $i=2,7$ | $i=1,6$ | |
| $\widetilde{y}_{i,s}^r$ (mm) | $0.17l_g$ | $0.34l_g$ | $0.5l_g$ | $0.66l_g$ | $0.83l_g$ | |

In order to realize the above modeling, a system consisting of 50 ordinary differential equations is used for describing the 3D structure dynamics of the vocal folds. In other words, for the motion of each mass element $m_{i,s}$ within the 3DM the ordinary differential equation is defined as

$$m_{i,s}\ddot{\mathbf{x}}_{i,s} = \vec{\mathbf{F}}_{i,s}^A + \vec{\mathbf{F}}_{i,s}^V + \vec{\mathbf{F}}_{i,s}^L + \vec{\mathbf{F}}_{i,s}^C + \vec{\mathbf{F}}_{i,s}^D,$$

$$\mathbf{x}_{i,s} = [x_{i,s}, \quad y_{i,s}, \quad z_{i,s}]^T. \tag{1}$$

$\mathbf{x}_{i,s}$ denotes the position of each mass element in the Cartesian coordinate system. $\dot{\mathbf{x}}_{i,s}$ and $\ddot{\mathbf{x}}_{i,s}$ denote the velocity and the acceleration, respectively. During phonation the acting forces which are considered in the 3DM are as follows:

(a) $\vec{\mathbf{F}}_{i,s}^A$: The anchor force acts as the function of the thyroarytenoid muscle in lateral direction,[45,52] see Eq. (A2).
(b) $\vec{\mathbf{F}}_{i,s}^V$: The vertical coupling force serves as internal force of the vocal fold tissue in vertical direction,[45,52] see Eq. (A4).
(c) $\vec{\mathbf{F}}_{i,s}^L$: Similarly, the longitudinal coupling force imitates the actions of the thyroarytenoid muscle and the vocal ligament in longitudinal direction, see Eq. (A5).
(d) $\vec{\mathbf{F}}_{i,s}^C$: During phonation the collision action between both vocal folds occurs, which causes the elastic structure deformation of the vocal folds, see Eq. (A8).
(e) $\vec{\mathbf{F}}_{i,s}^D$: Fluid driving force generated by glottal flow: As the driving origin of vocal fold dynamics, this force acts on the vocal fold surface from inferior to superior through the glottis, see Eqs. (A12)–(A14), (A20), and (A21).

All of the above-described forces act on the 3DM in three dimensions. The forces produced by damping elements are included in the different acting forces in detail, see Eqs. (A2), (A4), and (A5). The definitions and computation of these forces are based on the 2DM.[6,52] As an extension of the 2DM (Ref. 6) from 2D into 3D, the formulas of above-mentioned forces are illustrated in detail in the appendix with corresponding Figs. 12 and 13.

## B. Model parameter initialization

The 3DM is initialized to vibrate in accordance with the physiological characteristics of 3D vocal fold dynamics as

described in Sec. II A. Compared to the 2MM (Ref. 7) and 2DM (Ref. 6) the 3DM consists of high dimensionality on account of the spring stiffness ($50k_{i,s}^a$, $40k_{i,s}^l$, $40k_{i,s}^v$, $50k_{i,s}^c$), the mass $50m_{i,s}$, the damping coefficient ($50r_{i,s}^a$, $40r_{i,s}^l$, $40r_{i,s}^v$), the rest position $50\mathbf{x}_{i,s}^r$, as well as the pressure $5P_s$ yielding altogether 415 degrees of freedom.

As the initial configuration of the 3DM, a proper initialization for model parameters (i.e., stiffness, mass, thickness, and rest position) was chosen, as shown in Table I. It is based on the two parts: One is the 2MM presented by Ishizaka and Flanaga[7] and Schwarz et al.[2] The other is the diffusion properties of the vocal fold mucosal wave from inferior to superior as well as the phase delays obtained from excised larynx experiments,[26] see Table II. The initial stiffness $\widetilde{k}_{i,s}^l$ of longitudinal coupling spring is obtained by using Eq. (A6). The stiffness $\widetilde{k}_{i,s}^c$ is initialized after Eq. (A9). Moreover, in order to provide sufficient driving power for the 3DM, the initial subglottal pressure $\widetilde{P}^{sub}=2.4$ kPa has been proven to be adequate, which is located in normal range.[7,53] Additionally, within the 3DM, the glottal length $l_g$ is set to 0.7 cm for females and 1.4 cm for males, which situates within the range of literature values and is identical to the lengths chosen in former works.[1,2,11,12]

TABLE II. Different averages of 3D maximum displacements $(x,y,z)$, velocities $(v)$, accelerations $(a)$ and phase delays $(\phi_x,\phi_y,\phi_z)$ of 3D maximum displacements at six longitudinal lines (L1–L6, from inferior to superior) gathered from Boessenecker et al. (Ref. 26) $(x,y,z)$ denote the averages of maximum displacements in lateral, longitudinal, and vertical directions, respectively. Likewise, $(\phi_x,\phi_y,\phi_z)$ are for the phase delays in three directions, respectively.

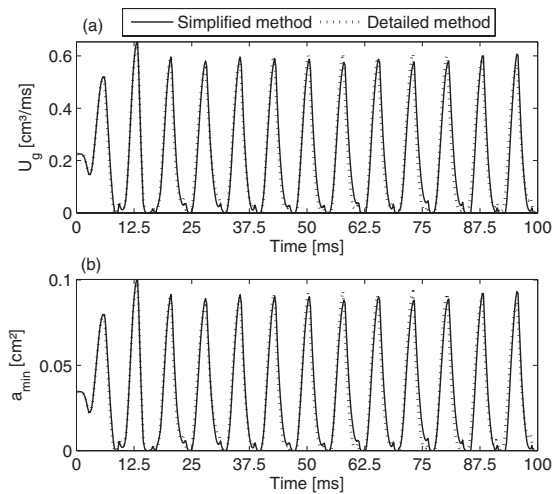| Suture lines | $(x,y,z)$ (mm) | $v$ (mm/ms) | $a$ (mm/ms) | $(\phi_x,\phi_y,\phi_z)$ (deg) |
|---|---|---|---|---|
| L6 | (1.29, 0.28, 0.91) | 0.86 | 0.21 | (170, 134, 51) |
| L5 | (1.57, 0.18, 0.95) | 0.98 | 0.24 | (160, 82, 24) |
| L4 | (1.52, 0.25, 0.5) | 0.87 | 0.22 | (84, 12, 17) |
| L3 | (1.09, 0.17, 0.47) | 0.59 | 0.15 | (39, −4, 10) |
| L2 | (0.64, 0.16, 0.37) | 0.37 | 0.08 | (18, −2, 6) |
| L1 | (0.39, 0.15, 0.33) | 0.29 | 0.08 | (9, 3, 8) |

FIG. 4. Upper figure: Glottal volume flow rate $U_g$ obtained by the simplified-method (solid line) and the detailed-method (dotted line). Lower figure: Minimum glottal area $a_{min}$ obtained with the simplified-method and the detailed-method.

## C. Pressure function

In order to compute the driving fluid forces, we discuss two methods to determine the pressure within the glottis. Based on the initial model parameters (i.e., standard model parameters, see Table I) the comparison of the volume flow rate $U_g$ generated from both methods is performed. By the simplified-method the calculation of the pressure function is in accordance to Steinecke et al.[8] and Schwarz et al.,[6] represented in Eqs. (A10)–(A14). However, by the detailed-method the approximate pressure distribution presented by Ishizaka et al.[7] with the jet function for a divergent glottis presented by Titze[48] is used to compute the pressures and corresponding driving forces at different planes, as shown in Eqs. (A15)–(A21).

## III. RESULTS

### A. Volume flow rate $U_g$ and minimal glottal area $a_{min}$

Demonstrating the applicability of the model, computational results for volume flow rate, minimal glottal area, as well as pressure and glottal area at each planes are presented. The applied parameters are based on initial values in Table I and with $k_{i,3}^v = 80$ N/m, $k_{i,3}^a = 80$ N/m, $k_{i,4}^a = 13$ N/m. Moreover, since the phonation threshold pressure (i.e., the minimum subglottal pressure required to initiate tissue oscillation, interpreted as a measure of ease of phonation)[24,54,55] is directly linearly related to the prephonatory glottal half-width, demonstrated in previous results and empirical observations.[23,24,55] To decrease phonation threshold pressure, the applied rest positions $\tilde{x}_{i,s}^r$ in the lateral direction are heuristically reduced to 32% of the initial values, so that the ease of phonation was resulted in. In Fig. 4(a) the volume flow rates, $U_g$ in the simulation of vocal fold oscillation using the 3DM with new adjusted model parameters by the simplified- and detailed-methods are shown. The duration of the simulation time was 100 ms, starting from rest state $t = 0$. Except the parts of launch phase of the simulation, the
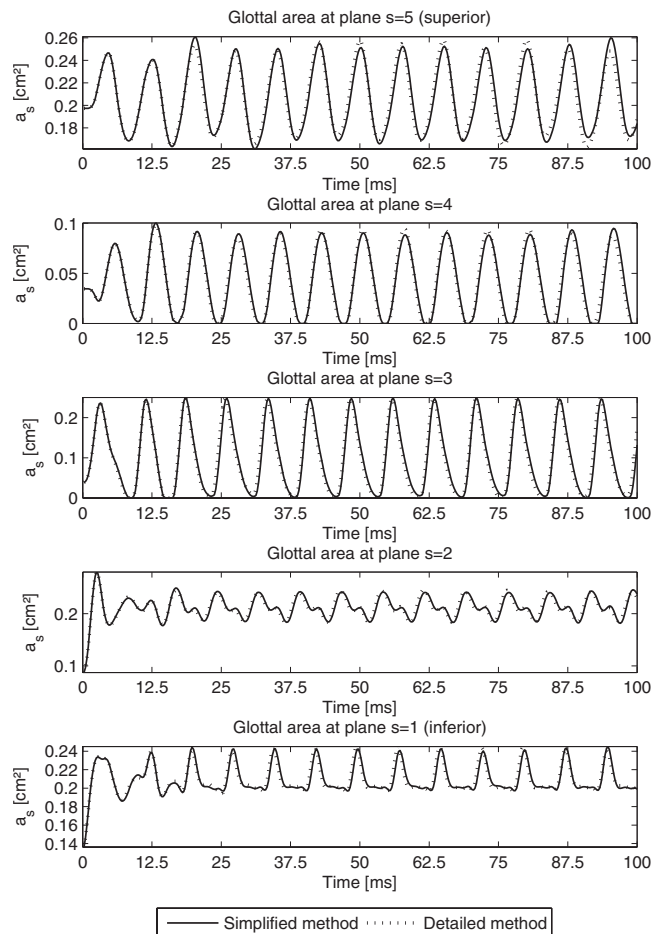


FIG. 5. Glottal area $a_s$ for each plane ($s = 1, \ldots, 5$) by using the simplified-method (solid lines) and the detailed-method (dotted lines).

maximum volume flow rates $U_g$ by using the simplified- and detailed-methods were, respectively, 0.607 and 0.609 cm³/ms.

The volume flow rate $U_g$ during the 3DM dynamics was periodically equal to zero due to sufficient glottal closure, see Fig. 4(a). The corresponding minimal glottal area $a_{min}$ is shown in Fig. 4(b). In 3DM the minimal glottal area occurred at plane $s = 4$ (corresponds to the edge of the vocal folds) and plane $s = 3$ (right below the vocal fold edge).

Tiny secondary peaks of minimal glottal area appeared, Fig. 4(b). These secondary peaks corresponded to alternating minimal values of glottal area at planes $s = 3, 4$ (Fig. 5). This means mass elements at plane $s = 4$ approached to glottal midline, meanwhile mass elements at plane $s = 3$ were still pushed away from the glottis center and vice versa. Hence, minimal glottal area phase delays between plane $s = 3$ and $s = 4$ created the small secondary peaks.

### B. Pressure $P_s$ and corresponding glottal area $a_s$

For optimizing the computation time of the 3DM, the detailed and simplified pressure distribution functions are compared. As shown in Fig. 6, pressures $P_s$ at different planes $s = 1, \ldots, 5$ within the 3DM by using the simplified- and detailed-methods are illustrated. The pressures at superior planes $s = 4, 5$ were derived zero. The shapes of the pres-
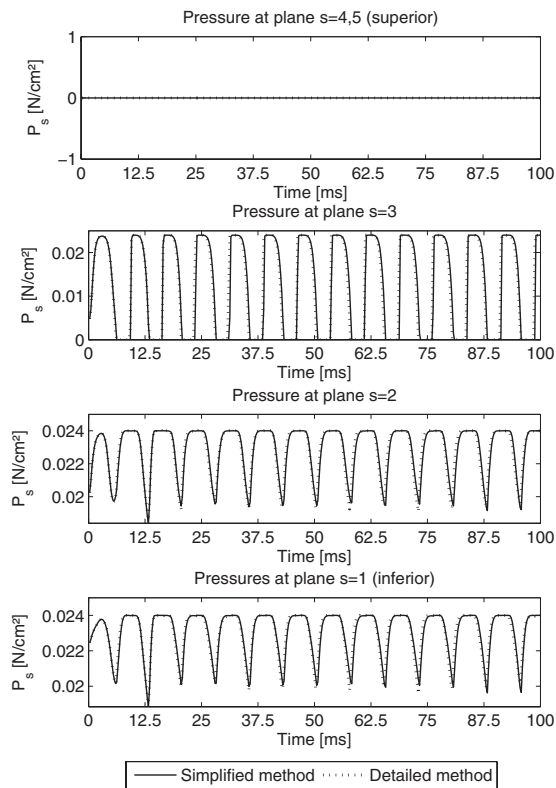
FIG. 6. Pressure $P_s$ for different planes ($s = 1, \ldots, 5$) by using the simplified-method (solid lines) and the detailed-method (dotted lines).

sures indicate that the pressure at third plane had strong fluctuations. The pressure at the most inferior plane $s = 1$ fluctuated the least. The corresponding frequencies for both pressure functions were 132 Hz and are in range of normal male voice. After the launch process of the 3DM, the maximum and minimum values of the pressures derived by the simplified- and the detailed-methods [Eqs. (A10)–(A21), see Appendix] are shown in Table III exhibiting similar values.

The corresponding glottal area $a_s$ at each plane $s$ by using the simplified- and detailed-methods are shown in Fig. 5. During stable vibration of the 3DM, the maximum and minimum values of the glottal area derived by the simplified- and detailed-methods can be seen in Table III. By means of the observation of variations for glottal area at each plane, we found that the amplitudes of the glottal area at planes $s$

$= 4, 5$ were larger than others. The glottal area at the most inferior plane was the smallest in accordance to the results in Boessenecker et al.[26] The glottal pressure distributions for both methods (simplified and detailed) were equal. Hence, from now on, the applied pressure distribution based on the simplified-method as described in Eqs. (A10)–(A14) is applied.

## C. Dynamics of the 3DM

In the following, important properties describing the structure dynamics of the 3DM are presented. The mucosal wave, the 3D trajectories, the displacements with corresponding fundamental frequencies, the velocities, as well as the accelerations are given for certain examples to show the model's applicability.

### 1. 3D movements and mucosal wave

To visualize the symmetric output dynamics generated with initial model parameters (Table I) within the 3DM, the dynamics of the mass elements situated on the right vocal fold during one oscillation cycle are displayed in Fig. 7. The model surface was interpolated using spline functions minimizing the flexion within curves.[56]

Correspondingly, movements of the mass elements on the right side during one oscillation cycle are described in Fig. 7 in detail. The dynamics of the 3DM start at rest position with rectangular glottis closure type; therefore, the trajectories for the launch phase of the 3DM were also displayed. The mucosal wave propagating on the vocal fold surface is clearly demonstrated. The corresponding duration of this cycle was 9 ms. In general, Fig. 7 emerges: In the beginning ($t = 0$) the 3DM was closed, while the intraglottal pressure increased. The mass elements located at plane $s = 3$ were still attached to the opposite side at midline. Due to the increase in intraglottal pressure, the mass elements located at planes $s = 1, 2$ on both sides began to deform measurably and were pushed away from midline. At $t = 2/7\pi$, the mass elements at plane $s = 3$ were pushed away from midline, while the mass elements at plane $s = 2$ continued being pushed away from midline. However, the mass elements at plane 4 already began to approach the midline. At $t = 4/7\pi$, the mass elements at plane $s = 4$ further approached the midline and moved upward. The mass elements at plane $s = 3$ were further

TABLE III. Maximum/minimum values of the pressures $P_s$ and the glottal area $a_s$ at each plane by using the simplified- and detailed-methods after the launch phase of the model.

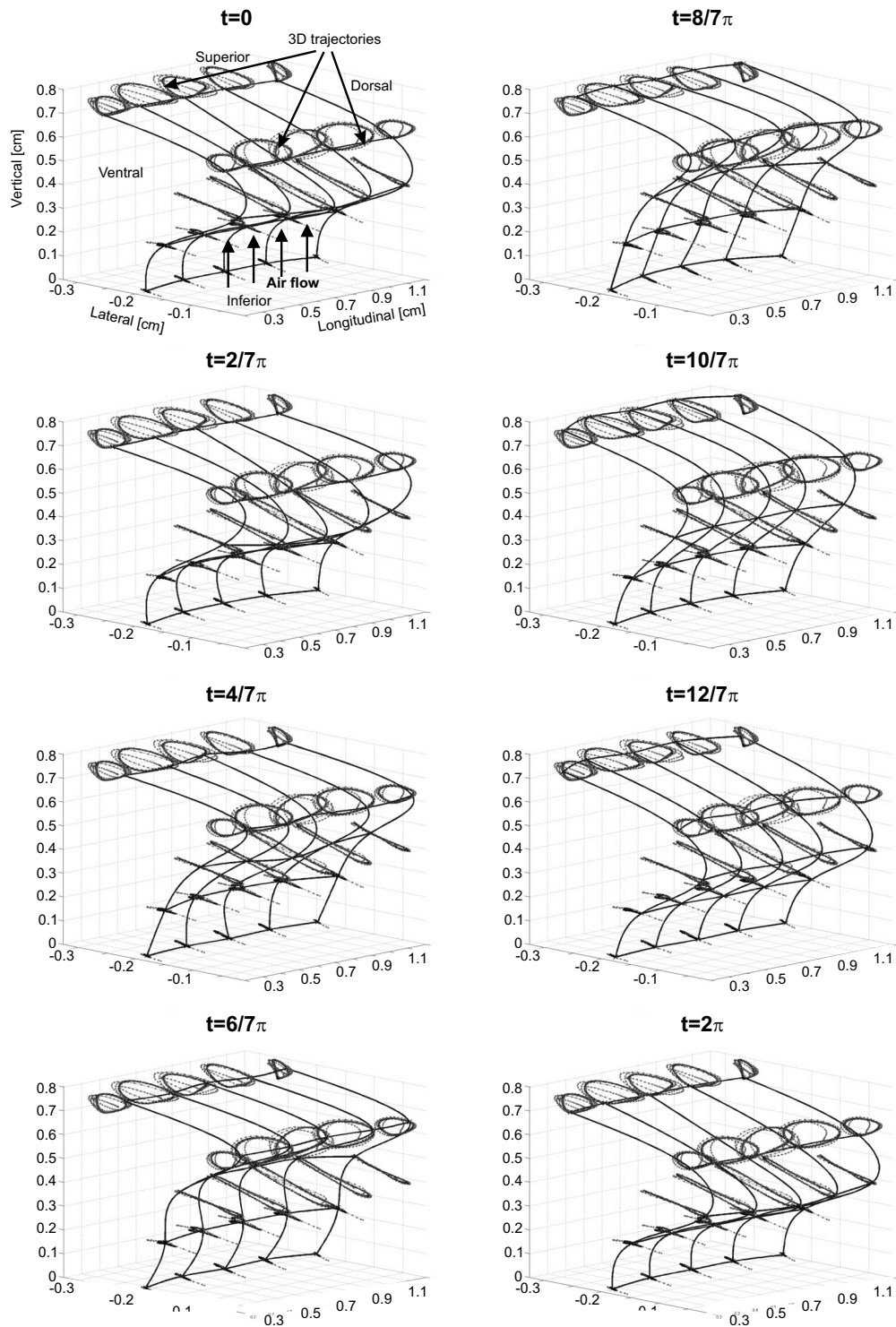| | Plane $s$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| | | | | Pressure | | |
| Simplified | Max. $P_s$ (N/cm²) | 0.024 | 0.024 | 0.024 | 0.0 | 0.0 |
| | Min. $P_s$ (N/cm²) | 0.02 | 0.0192 | 0.0 | 0.0 | 0.0 |
| Detailed | Max. $P_s$ (N/cm²) | 0.024 | 0.024 | 0.024 | 0.0 | 0.0 |
| | Min. $P_s$ (N/cm²) | 0.0196 | 0.0191 | 0.0 | 0.0 | 0.0 |
| | | | | Glottal area | | |
| Simplified | Max. $a_s$ (cm²) | 0.2434 | 0.2457 | 0.25 | ?0,0934 | 0.2578 |
| | Min. $a_s$ (cm²) | 0.1982 | 0.1824 | 0.0 | 0.0 | 0.1673 |
| Detailed | Max. $a_s$ (cm²) | 0.2447 | 0.2452 | 0.2480 | 0.0949 | 0.2582 |
| | Min. $a_s$ (cm²) | 0.1973 | 0.1844 | 0.0 | 0.0 | 0.1647 |

FIG. 7. Sequence of 3D dynamics for all mass elements located at one vocal fold during one oscillation cycle, with a spline-interpolated 3DM surface. The dotted lines represent the trajectories of each mass element.

pushed away from midline. The superior mass elements at plane $s=5$ were elevated. From time step $t=2/7\pi$ to $t=6/7\pi$, the mass elements at upper planes ($s=3,4,5$) moved upward due to high intraglottal pressure. From time step $t=8/7\pi$ to $t=12/7\pi$, the corresponding mass elements sank because of relaxation under the low intraglottal pressure. At $t=12/7\pi$, the model began to close, meanwhile the increase in intraglottal pressure started again. At $t=2\pi$, the model was

closed again and the intraglottal pressure was gradually established. Glottis closure occurred at t planes $s=3,4$.

In order to inspect the quality of the 3DM, a comparison between the 3D trajectories generated from 3DM and the 3D trajectories extracted from excised human hemilarynges[26] at middle cross-section is depicted in Fig. 8. Owing to the tiny movements of the micro-suture located at the most inferior sagittal line [Fig. 8(b)], the dynamics at these micro-sutures

Yang *et al.*: Modeling of three-dimensional vocal fold dynamics
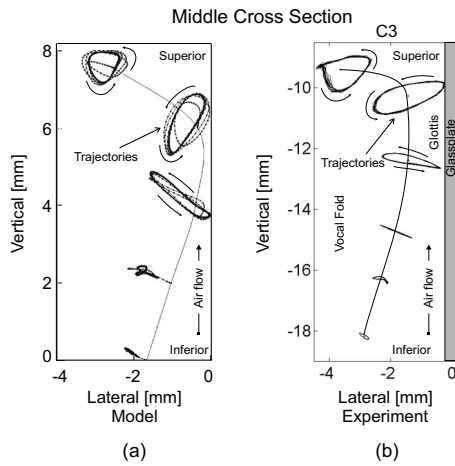
FIG. 8. (a) Five simulated 3D trajectories (dotted lines) along middle vertical cross section produced by 3DM. (b) Six experimental 3D trajectories (dotted lines) along middle vertical cross section extracted from an excised human hemilarynx (Ref. 26). The vertical contours in both images correspond to the rest positions of the mass elements as well as the vocal fold tissue, respectively. For all five mass elements (from inferior up to superior) in (a) the movement ranges can be roughly described as 0.5, 1, 2, 2, and 1.5 mm, respectively. By comparison, the movement ranges of the upper five suture points (from inferior up to the vocal fold edge) in (b) are almost 0.3, 1, 2, 2, and 1.7 mm, respectively. Hence, the trajectories show similar behavior.
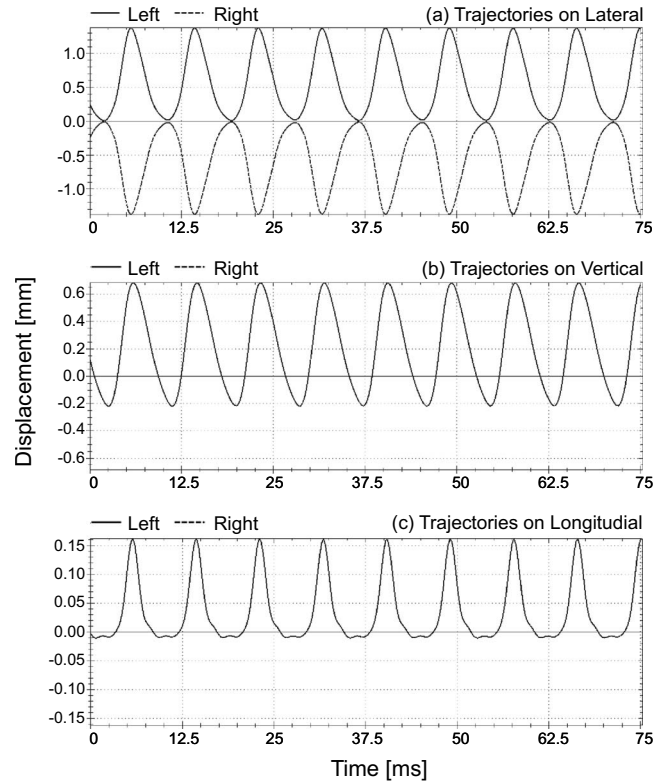


FIG. 9. Simulated symmetrical trajectories in lateral direction (upper figure), vertical direction (middle figure), as well as longitudinal direction (lower figure) of two opposing mass elements (dotted lines for the mass element $m_{2,3}$ on right side, solid lines for the mass element $m_{7,3}$ on left) which are situated at third plane and second cross-section in the symmetrical 3DM.

were neglected in 3DM. This means that only the movements of the micro-sutures located at the upper five sagittal lines (five planes) were taken into account. The shape of movement of the mass element at fifth plane (i.e., the most superior plane) was like a triangle, being similar to the experimental movement of the suture point with uppermost position of the excised vocal fold surface. In the fourth plane (i.e., around the vocal fold edge) the movement of the mass element in 3DM was much like an ellipse, which was similar to the experimental dynamics of the suture point at the corresponding plane, see Fig. 8. In third plane (i.e., right below the vocal fold edge) the movement of the mass element was like a flat ellipse which was similar to the experimental dynamics of the suture point at L4 (fourth sagittal line). The movement shapes of mass elements in lower planes ($s=1,2$) were like line segments, which were similar to the experimental movements of micro-suture at L2 and L3. Both models show a significant decrease in movements from medial to inferior.

### 2. 3D trajectories

Standard model parameters (Table I) were chosen now. The corresponding glottis closure type was rectangle. In order to show the model application, the absolute values $|\tilde{x}_{i,s}^r|$ for lateral rest positions of the masses from inferior to superior were temporarily reset to 0.988, 0.634, 0.280, 0.251, 1.431 mm (i.e., convergent glottis). In Fig. 9 this example of 3D trajectories extracted from the symmetric 3DM is described. These 3D trajectories corresponded to mass elements $m_{2,3}$ and $m_{7,3}$ on both sides, respectively. Their corresponding movements in lateral direction were exactly symmetrical, see Fig. 9(a). Due to the symmetric characteristics of the healthy vocal folds as assumed in this work, the

vertical and the longitudinal movements for mass elements which were located on both sides (left and right) at the same plane and same cross-section were identical, see Figs. 9(b) and 9(c).

### 3. 3D displacement maxima

Figure 10 shows the 3D displacements of mass elements within the 3DM, being generated by the symmetric initial model parameters (from now on, the rest positions for mass elements were defined as shown in Table I again). Due to symmetry, only one vocal fold side was depicted. The displacements of mass elements were symmetric, respectively, in the longitudinal direction and centered on the middle cross-section. The applied glottis closure type was rectangular. The maximum displacements of the mass elements can be described as follows.

*a. Lateral displacement maxima.* Figure 10(a) shows the comparison of the maximum lateral displacements belonging to the different mass elements $m_{i,s}$. The range of the maximum displacements in lateral direction was between 0.09 mm ($m_{1,1}$ and $m_{5,1}$) and 1.47 mm ($m_{3,3}$). This means that the highest lateral displacement occurred at plane $s=3$ on middle cross-section. Overall compared to the corresponding longitudinal displacements, the lateral displacement maxima were much higher. Additionally, for all mass elements the lateral displacements increased from the most inferior plane ($s=1$) up to the medial plane ($s=3$). Similarly,
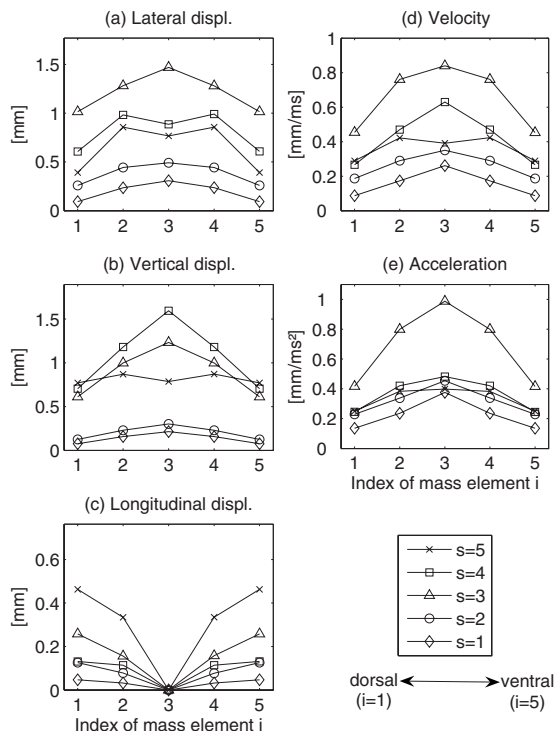
FIG. 10. [(a)–(c)] Maximum lateral, vertical, and longitudinal displacements of each mass element on one side of the 3DM. (d) Absolute maximum velocities of each mass elements on one side of the 3DM. (e) Absolute maximum accelerations of each mass elements on one side of the 3DM. $s = 1, \ldots, 5$ denote the planes from inferior up to superior.

an increase from the most superior plane ($s=5$) to the medial plane ($s=3$) occurred. Along longitudinal direction the lateral displacements decreased from middle cross-section to dorsal/ventral cross-section, except mass elements $m_{3,4}$ and $m_{3,5}$.

*b. Vertical displacement maxima.* The vertical displacement maxima ranged within 0.08 and 1.6 mm, Fig. 10(b). The highest vertical displacement maximum was situated at mass element $m_{3,4}$, which was located in middle cross-section at plane $s=4$. The second highest vertical displacement maximum occurred at medial plane ($s=3$) on the middle cross-section. The lowest vertical displacement maxima were at the ventral and dorsal mass elements ($i=1,5$) at the most inferior plane ($s=1$). The vertical displacement maxima at superior planes ($s=3,4,5$) were nearly identical on the ventral and dorsal cross-sections ($i=1,5$). At inferior planes ($s=1,2$) the vertical maximal displacement values were similar and significantly decreased compared to planes $s=3,4,5$. In general, the vertical displacement maxima increased from inferior up to superior and decreased from middle cross-section to dorsal/ventral cross-section.

*c. Longitudinal displacement maxima.* The appearance of longitudinal displacements is caused by longitudinal coupling forces with corresponding stiffnesses as described in Eqs. (A5)–(A7). Figure 10(c) shows the computed maximal longitudinal displacement values. The longitudinal displacement maxima ranged between 0 and 0.46 mm. The smallest longitudinal displacement occurred at the most inferior plane ($s=1$). The largest longitudinal displacement was found at the most superior plane ($s=5$). Compared with the corre-

sponding lateral and vertical displacement maxima, the longitudinal displacement maxima were the smallest. Due to longitudinal symmetry of the initial model parameters, the longitudinal displacement maxima were equal to zero at middle cross-section. For all five planes, the longitudinal displacement maxima decreased from ventral ($i=5$) and dorsal ($i=1$) cross-sections to middle cross-section ($i=3$).

*4. Velocity and acceleration maxima of dynamics*

In order to get motion characteristics of the 3DM, the corresponding distributions of velocity and corresponding acceleration maxima values are shown in Figs. 10(d) and 10(e).

*a. Velocity maxima.* The maximal computed velocity values are shown in Fig. 10(d). Velocity maxima ranged from 0.09 up to 0.84 mm/ms. The highest velocity maximum was located on mass element $m_{3,3}$. The lowest velocity maxima were found on mass elements $m_{1,1}$ and $m_{5,1}$. In general, for all mass elements the velocity maxima increased from ventral/dorsal cross-section ($i=1,5$) up to middle cross-section ($i=3$), except mass element $m_{3,5}$. For all five cross-sections ($i=1,\ldots,5$), the velocity maxima increased from the most inferior plane ($s=1$) to the medial plane ($s=3$). Likewise, from the most superior plane ($s=5$) to the medial plane, there was also a significant increase in velocity maxima, except dorsal/ventral cross-section at plane $s=5$.

*b. Acceleration maxima.* Figure 10(e) shows the maximal computed accelerations. The maximal acceleration values ranged between 0.14 and 0.99 mm/ms$^2$. It was quite obvious that the acceleration maximum at medial plane ($s=3$) were the highest, especially at middle cross-section ($i=3$). The lowest acceleration maxima were at the most inferior plane ($s=1$), especially at ventral/dorsal cross-section ($i=1,5$). For other planes (except planes $s=1,3$), the acceleration maxima were similar and decreased significantly compared to those at plane $s=3$. For all mass elements the acceleration maxima increased from dorsal/ventral cross-section to middle cross-section.

**D. Comparison of dynamics**

In order to demonstrate the applicability and validation of the proposed model, the 3D structure dynamics are generated by the 3DM with different subglottal pressures $P^{\mathrm{sub}}$. Since the range of the subglottal pressure is mostly described from 0 to 3.5 kPa,[7,57] we selected the subglottal pressures $P^{\mathrm{sub}}$ as 1.2, 1.6, 2, 2.4, 2.8, and 3 kPa. The longitudinal length $l_g$ was set to 14 mm (male). For all pressure values, inital model parameters were chosen, Table I. The glottis closure type is always set as rectangular. Altogether six examples will be discussed.

Due to longitudinal and lateral symmetrical definition of the model parameters, the comparison of the six examples is performed for one vocal fold side along the vertical middle cross-section in the 3DM. Figure 11 shows the displacement, the velocity, and the acceleration maxima of the six examples. Owing to symmetries of predefinition for the examples, the maximal longitudinal displacement values for mass elements on the middle cross-section were derived
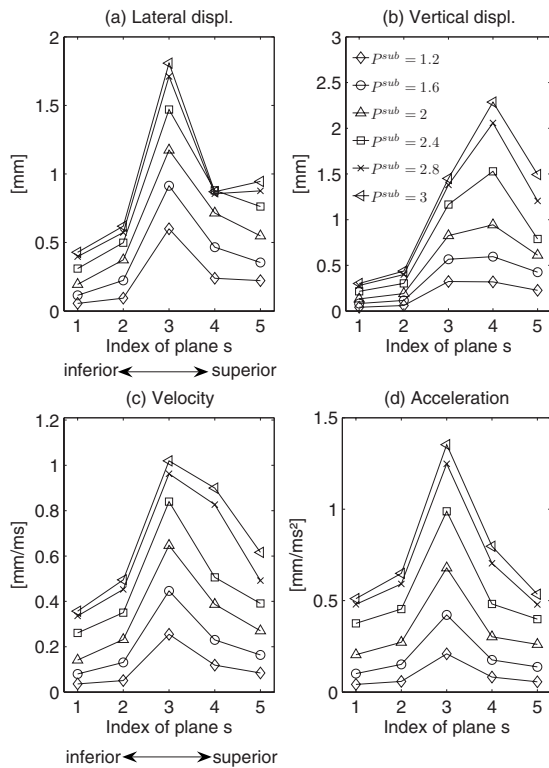
FIG. 11. Displacement, velocity and acceleration maxima of mass elements on the right side of the middle cross-section for different subglottal pressure configurations. The maximum displacement values are illustrated in lateral and vertical directions.

zero. In general the higher the subglottal pressure was set to, the greater the displacement and velocity as well as acceleration maxima were achieved. The fundamental frequencies by different subglottal pressures (1.2, 1.6, 2, 2.4, 2.8, and 3 kPa) were derived as 104, 108, 108, 108, 108, and 108 Hz.

### 1. Comparison of displacement maxima

Figure 11(a) shows the computed maximal lateral displacement values for the six different subglottal pressures. The range of maximal lateral displacement values was between 0.06 mm ($P^{sub}$=1.2 kPa, $s$=1) and 1.81 mm ($P^{sub}$=3 kPa, $s$=3). For all mass elements on middle cross-section, an increase in subglottal pressure resulted in a significant increase in lateral displacements. This increase of lateral displacement maxima was more explicit for the three upper planes ($s$=3,4,5), except for subglottal pressure $P^{sub}$=2.8, 3 kPa.

Figure 11(b) shows the maximal vertical displacement values by different subglottal pressures. Vertical displacement maxima ranged from 0.04 mm ($P^{sub}$=1.2 kPa, $s$=1) up to 2.29 mm ($P^{sub}$=3 kPa, $s$=4). As found for lateral displacement directions, an increase in subglottal pressure yielded a marked increase in vertical displacements for all mass elements in middle cross-section. Likewise, this increase in vertical displacement maxima was higher for the three upper planes ($s$=3,4,5).

### 2. Comparison of velocity and acceleration maxima

The maximal velocity values for different subglottal pressures are shown in Fig. 11(c). The range of maximal velocity values was between 0.03 mm/ms ($P^{sub}$=1.2 kPa, $s$=1) and 0.99 mm/ms ($P^{sub}$=3 kPa, $s$=3). An increase in subglottal pressure also yielded larger velocity values.

Figure 11(d) displays the computed maximal acceleration values for all mass elements at middle cross-section of one side. The range of values was from 0.02 mm/ms ($P^{sub}$=1.2 kPa, $s$=1) up to 1.44 mm/ms ($P^{sub}$=3 kPa, $s$=3). As found for velocity maxima of mass elements by different subglottal pressures, an increase in subglottal pressure resulted in a significant increase in acceleration values.

## IV. DISCUSSION

In clinical practice, quantitative analysis of the data captured by HS digital imaging can be a daunting task. It can be realized by directly describing and analyzing laryngeal characteristics such as fundamental frequency, vocal fold length, and vocal fold displacements.[41,42,58,59] Using the instantaneous frequency and amplitude obtained from high-speed laryngeal imaging recordings, the glottal perturbation can be measured.[60] Also investigated were growth of vocal fold amplitudes, subglottal pressure as well as myoelastic forces reflecting the phonation onset dynamics.[61–63]

For further supporting clinical diagnosis of voice disorders, objective quantification of the vocal fold 3D dynamics is meaningful. In order to visualize and objectively quantify spatiotemporal characteristics in vocal fold 3D dynamics with reasonable computational costs, a new 3D biomechanical modeling approach was described. Therefore, by means of the 3DM (Fig. 2) as extension of the 2DM,[2] the 3D dynamics of the entire vocal folds are rationally reproduced. Through the modification of model parameters describing physiological properties (i.e., mass and stiffness distributed in the entire larynx), the model simulates vocal fold dynamics by using 25 control points distributed along the surface of each vocal fold.

The model parameters are appropriately initialized to comply with three aspects: (1) The model generated 3D dynamics are similar to 3D dynamics extracted from hemilarynx experiments,[21,26–28,32] (2) similar to synthetic larynx models[36,37] and finite-element models,[38,39] and (3) the 3DM inherits the biomechanical numerical characteristics of vocal fold dynamics, which have been proposed and justified in earlier works.[7,45,46,64] Its model parameters have the values which match the vocal fold physiology.

### A. Generation of the volume flow rate, pressures, and glottal area

In this work, the volume flow rate $U_g$ during the 3DM dynamics using the initial model parameters shown in Table I after phonation onset is periodically equal to zero. This result [see Fig. 4(a)] is in agreement with the volume flow rate presented in Refs. 65 and 66 as well as Ref. 67. In accordance with Eq. (A18), it can be explained that the minimum glottis area $a_{min}$ in our case periodically amounts to zero [see Fig. 4(b)] due to regularly complete glottal closure
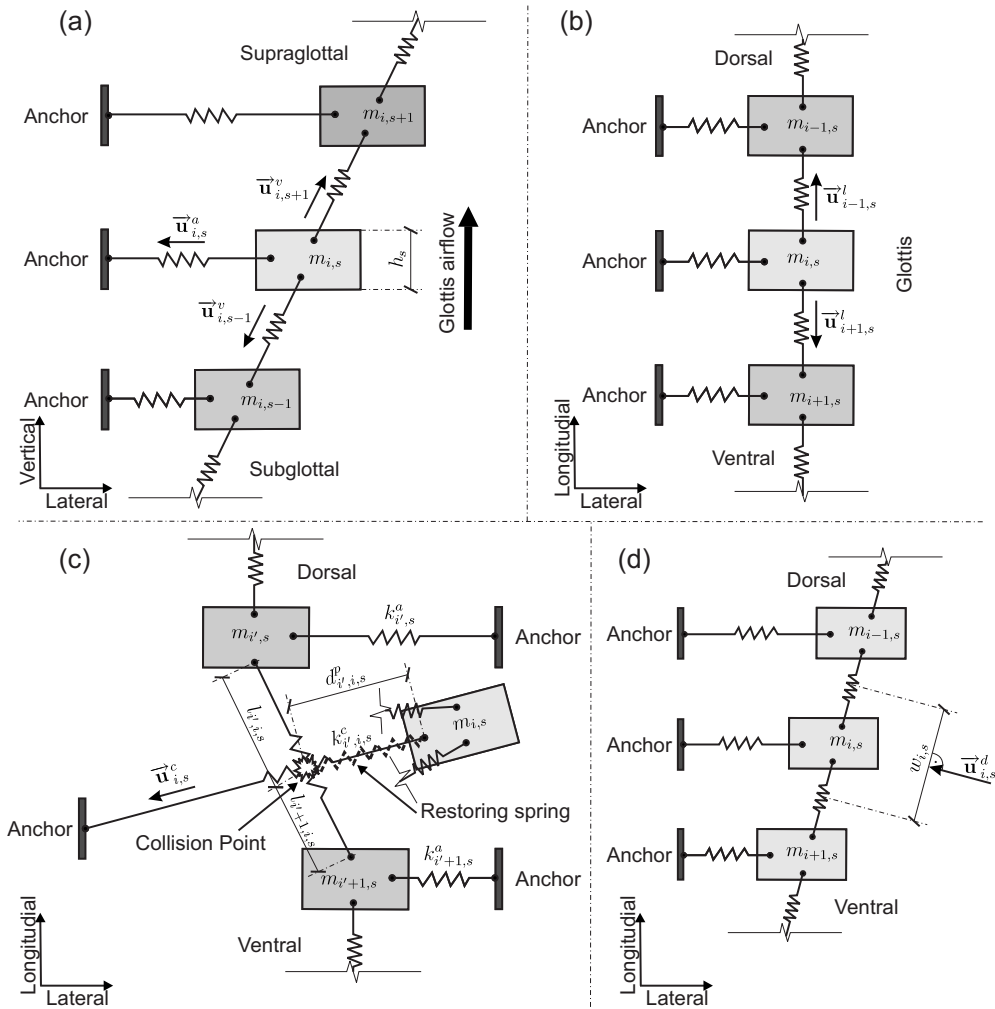
FIG. 12. (a) Graphical definition of the unit vector $\vec{\mathbf{u}}_{i,s}^a$ for mass element $m_{i,s}$ in the computation of the anchor force. Unit vector $\vec{\mathbf{u}}_{i,s-1}^v$, $\vec{\mathbf{u}}_{i,s+1}^v$ aiming at mass element $m_{i,s}$ for the vertical coupling force. $h_s$ means thickness of the mass element $m_{i,s}$. (b) Unit vector $\vec{\mathbf{u}}_{i-1,s}^l$, $\vec{\mathbf{u}}_{i+1,s}^l$ for mass element $m_{i,s}$ in the computation of longitudinal coupling force. (c) Graphical definition of collision and corresponding unit vector $\vec{\mathbf{u}}_{i,s}^c$ for mass element $m_{i,s}$ in the calculation of collision force. When the collision on mass element $m_{i,s}$ occurs, the anchor spring of mass element $m_{i,s}$ crosses through the longitudinal coupling spring between two adjacent mass elements $m_{i',s}$ and $m_{i'+1,s}$. $k_{i',i,s}^c$ indicates the equivalent stiffness of the restoring spring (Ref. 6) $d_{i',i,s}^p$ denotes the penetration depth during collision. (d) Graphical definition of the effective side length $w_{i,s}$ and the unit vector $\vec{\mathbf{u}}_{i,s}^d$ for the mass element $m_{i,s}$ in the computation of the driving force.

during vocal fold oscillations. It conforms to the phenomenon observed in the hemilarynx experiments.[26,30] In addition, glottal closure insufficiency can also occur in normal phonation, especially for women in the dorsal region.[42] However, the common case for normal phonation goes along with complete glottal closure.

The obtained pressures in the 3DM are constantly equal to zero for $s=4,5$. For $s=3$ the pressure is periodically equal to zero as presented by Lowell *et al.*[65] and Jiang *et al.*[68] as well as Story *et al.*[47] However, they derived also negative intraglottal pressure values, which is excluded due to Eqs. (A10) and (A19) by definition in our model. For $s=1,2$ the pressure is almost periodically zero. This can be explained by assumptions in Refs. 8 and 48 as well as Ref. 69. The net intraglottal pressure is positive in the case of the convergent orifice (since it has to decrease to a value of zero) and negative in the case of the divergent orifice (it has to increase to a value of zero). The airstream often detaches from the surface of the vocal folds and forms a jet[48] which keeps the pressure in the glottis closer to zero. Bernoulli's law is less
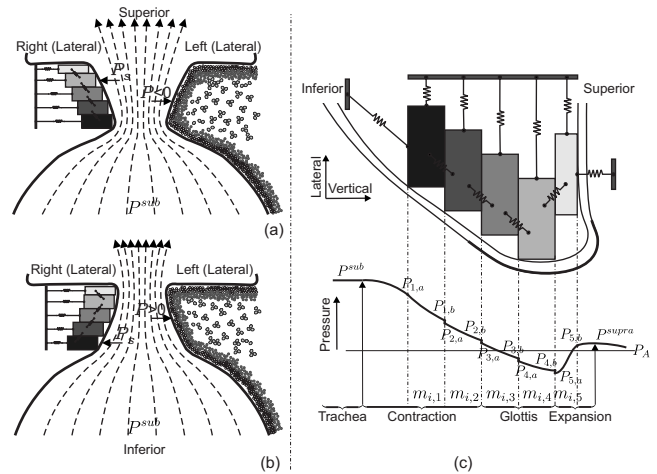


FIG. 13. (a) Airflow and driving pressure in the divergent glottis, including the cross-sectional structure of the 3DM. (b) Airflow and driving pressure in the convergent glottis where the driving pressure $P$ acts on the tissue. Accordingly, in the 3DM, $P_s$ acts on mass elements of each plane ($s=1,\dots,5$). (c) Detailed pressure computation. $P_A$ denotes the atmospheric pressure (Ref. 66). According to the assumption from Titze (Ref. 48) and conclusion in Ref. 69. $P_A$ as well as $P^{supra}$ is set to zero.

applicable to the divergent glottis. This assumption applied in Eqs. (A10) and (A20) yielded that pressures at upper planes $s=4,5$ were derived zero. This result supports the conclusion demonstrated in Spencer et al.:[36] *Flow studies indicate that fluid pressure acts mainly on the inferior surface.*

Moreover, secondary peaks of minimal glottal area illustrated in Fig. 4(b) explicitly demonstrate that the minimal values of glottal area alternatively arise at different planes ($s=3,4$, see Fig. 5). Similarly, the secondary peaks also occurred in the envelope of volume flow rate [Fig. 4(a)] dependent on $a_{min}$ as found in Ref. 70. This phenomenon is in agreement with conclusions about glottal area waveforms presented in Koizumi et al.[66] and Wong et al.[52] as well as Ishizaka et al.[7] It also reveals that phase difference of glottal areas between different planes exists,[7] which is due to the displacement phase delays occurring during mucosal wave propagation as concluded in Ref. 26. However, no phase delays existed in pressures $P_s$ ($s=1,\dots,5$) due to the intraglottal pressure following the Bernoulli equation, see Fig. 6.

For the 3DM two pressure functions were compared, i.e., a simplified- and a detailed-method. The simplified-method originates from Schwarz et al.[6] and Steinecke et al.[8] In the detailed-method, pressure distribution along the glottal flow is taken into account. The main difference between both methods is the calculation of the glottal pressure. As a result derived from Fig. 4, the pressures, glottal area, and corresponding volume flow rate as well as minimum glottal area produced from both methods were identical. This means that the simplified-method yields similar pressure distribution as produced by the detailed-method. The reason that little difference was found in the results for the simplified- and detailed-methods might be that both of them assume no glottal viscous losses. It is consistent with the assumption that glottal viscous losses is neglected in the model. Due to simplicity of computation of the simplified-method in contrast to the detailed-method, we used and suggested the simplified-method.

## B. Overall dynamics of the 3DM

In contrast to prior works,[7,45,46,64] the 3D model dynamics included the vertical component and simulated the mucosal wave, see Fig. 7. During the oscillation cycle the mass elements located at the superior planes $s=4,5$ moved generally lateral and downward during the open period, upward during the closed period, so as to finish the previous oscillation cycle. The movements of the mass elements at medial plane $s=3$ and inferior planes $s=1,2$ were predominantly performed in lateral direction (Fig. 7). These resulting vertical dynamics are comparable to experimental results from excised human vocal folds[26] and *in vivo* canine vocal folds.[30]

Additionally, Fig. 7 implicitly supports the conclusions presented by Döllinger et al.[19] and Berry et al.[32] that a low-pressure condition is created by a divergent glottis (supports the folds to approximate) and a high-pressure condition is created by a convergent glottis (support to push the vocal folds apart), see Fig. 8 in Ref. 32 and Fig. 3 in Ref. 19.

### 1. Mucosal wave propagation

In prior studies, with the aids of hemilarynx data[21,26] as well as experimental data from excised canine larynges[30–32] mucosal wave propagation in normal phonation was analyzed and quantified. It was confirmed that lateral phase delays were much higher than vertical or longitudinal phase delays. Results in Ref. 26 gave lead for displacement maxima in 3D: $0.26-1.90$ mm (lateral), $0.21-1.15$ mm (vertical), and $0.10-0.33$ mm (longitudinal). Similarly, such phenomenon including vertical and lateral vibrations was reported in a study on aerodynamically and acoustically driven modes of the vocal fold vibration by using a physical model.[35]

In this work, the model dynamics are displayed along the vertical direction from inferior to superior through the entire glottis, see Fig. 8(a), which reflects the mucosal wave propagation of the human vocal folds. It shows the simulated 3D trajectories of each mass element at five planes being similar to the corresponding 3D trajectories extracted from previous experiments of excised vocal folds.[19,26] The simulated 3D trajectories of the mass element located at fifth plane, i.e., the most superior plane in the vertical direction, is like a triangle. Meanwhile, the experimental movement of the micro-sutures with the uppermost position at the surface of excised vocal folds is very much alike. At the fourth plane, i.e., around the vocal fold edge, the movement of the mass element is nearly elliptical, which is similar to the experimental dynamics of the micro-sutures at the corresponding plane. It moved upward and laterally to reach the most upper lateral position, then moved downward to reach the lowest position on the way back.[34] In fact, the trajectories of mass elements at the fourth plane suggest that the maximum vertical displacement corresponds to closed position and the minimum vertical displacement corresponds to maximum open position, which is in conformity with the observations in study of HS calibrated 3D imaging of human vocal fold vibration dynamics by George et al.[71] This shows that the 3D model reproduces successfully experimental investigated human 3D dynamics. It is also in agreement with results obtained from excised canine larynx experiments,[32] which revealed that the elliptical paths traversed by the fleshpoints of the left vocal fold over 1 cycle of vibration increased from inferior to superior (see Fig. 5 in Ref. 32). It supports the conclusion from Döllinger et al.[30] that during normal phonation mucosal wave propagated primarily from inferior to superior. Similar to 3D movements extracted from finite-element models,[38,39] this result can be used as an evidence to demonstrate the existence of increase of elliptical excursions along the vocal fold surface from inferior to superior. From a different perspective, the mucosal wave propagation might actually diminish the effects of impact force during the closing phase of the glottis.[40]

Additionally, the movements of mass elements within the 3DM are in the same range as the recorded displacements of human vocal folds in Baer[72] Specifically, the movement range of particle 3 (at the subglottal surface, approximately, 5 mm below the supraglottal surface, see Fig. 3(b) in Ref. 72) is about 1 mm. In the 3DM as well as the hemilarynx experiment[26] the lower mass element [at the plane $s=2$, ap-

proximately 5 mm below the most superior mass element, see Fig. 8(a)] and the lower suture point [at the layer L3, approximately 5 mm below the supraglottal surface in the hemilarynx experiments, see Fig. 8(b)] have the same movement range about 1 mm. The results can be used to reflect the rationality of the proposed 3DM relative to laryngeal experiments. Also, it confirmed the validity of earlier vocal fold physiological study in Baer.[72]

### 2. Displacement, velocity, and acceleration maxima during dynamics

In prior studies,[36,37,69,73–75] to determine mechanical fields on the superior surface of vocal folds as well as physical processes synthetic larynx models were proposed. As found in experiments of human hemilarynges, the largest inferior-superior vibratory displacements occurred in the center of the model fold [Figs. 4(a) and 4(b) in Ref. 36]. Additionally, Fig. 6 in Ref. 38 exhibited that the highest excursion of the vocal folds was at the upper parts.

In this work, the displacement maxima between the different mass elements were described in lateral, longitudinal, and vertical directions, respectively. Overall, the longitudinal displacement maxima were lower than the corresponding lateral and vertical components. The highest displacement maxima were mostly located at medial and superior planes (i.e., $s=3,4,5$). For each direction the highest values were, respectively, lateral 1.47 mm ($m_{3,3}$), longitudinal 0.46 mm ($m_{1,5}, m_{5,5}$), and vertical 1.6 mm ($m_{3,4}$). However, results obtained from human hemilarynx experiments in Refs. 26 and 19 and as well as excised canine larynx experiments in Refs. 30 and 32 and study of *in vivo* canine vocal fold dynamics in Ref. 30 indicate that lateral displacement maxima (medial-lateral) were the highest, and the longitudinal displacement maxima (anterior-posterior) were the smallest. Furthermore, previous work using a synthetic larynx model presented in Ref. 36 and excised canine larynx experiments in Ref. 34 show that vibratory displacements in the medial-lateral direction are significantly larger than in the vertical orientation (inferior-superior), see Figs. 4(c) and 4(d) in Ref. 36. Although similar to studies of finite-element models of vocal fold vibration,[38,39] in this work the vertical displacement maxima were higher than the displacement maxima in lateral and longitudinal directions. However, the vertical displacement maxima were only slightly higher than the displacement maxima in lateral direction.

Moreover, Figs. 10(a)–10(c) reveal that the largest elliptical dynamics are near the superior medial vocal fold and decrease within the inferior part. This result is in agreement with corresponding conclusions obtained from hemilarynx experiments in Ref. 19, determination of mechanical fields using a synthetic larynx model in Ref. 36, and a self-oscillating finite-element model in Ref. 76 as well as excised canine larynx experiments in Ref. 32. It can be explained by the combined effect of tissue deformation and airflow pressure generating a higher amplitude vibration than for the inferior part.

According to previous studies,[19,26,30,31] the velocity and acceleration maxima of the mucosal movements increased from inferior to superior. In our work, Fig. 10 reveals that the largest velocity and acceleration maxima were located at planes (i.e., $s=3,4$) which correspond to the region around the human vocal fold edge (Fig. 7). This result might be used to implicitly support the suggestion proposed by Spencer *et al.*[36] that *the medial surface is a free surface before collision onset, the stresses (depend on intraglottal pressure) orthogonal to that surface are small during the open phase.*

### C. Applicability of the 3DM

The applicability of the 3DM was shown for six different subglottal pressure conditions. In order to demonstrate the model's applicability to various subglottal pressures $P^{sub}$, we made the emphasis on the comparison of the dynamic characteristics (i.e., displacement, velocity, and acceleration maxima). The discussed dynamic characteristics belong to mass elements located on the middle cross-section lying on the axis of the longitudinal symmetric 3DM predefined with the standard model parameters.

In the comparison of model dynamics in both directions (lateral/vertical) under different subglottal pressures, Figs. 11(a) and 11(b) reveal that larger subglottal pressures yield higher lateral and vertical displacements, especially within superior medial and superior part of the vocal fold. In accordance with Fig. 6 presented in Ref. 76 the vibration amplitude of the vocal fold is roughly the square root of the subglottal pressure, which was found by Titze.[77] This can be used to explain the higher displacements by increasing subglottal pressure.

The velocity and acceleration maxima are also significantly higher with larger subglottal pressure [Figs. 11(c) and 11(d)]. This result is in agreement with Spencer *et al.*,[36] which confirms that velocity is larger for higher subglottal pressure than that for lower subglottal pressure. For acceleration, a study by Jiang *et al.*[78] on canine larynges supports our findings.

Additionally, results in Fig. 11 associated with the behavior of fundamental frequencies imply that mucosal wave velocity $v$ had no relation to fundamental frequency, unlike results obtained by Titze *et al.*[31,79] that mucosal wave velocity increases linearly with fundamental frequency. Moreover, it was found in Ref. 36 that mucosal wave velocity increased as subglottal pressure increased, while fundamental frequency was regarded as a function of subglottal pressure.[7] However, in this work the fundamental frequencies by different subglottal pressures were identical (108 Hz), except that at $P^{sub}=1.2$ kPa it was slightly smaller (104 Hz). Compared to the behaviors for $\eta=0$ in Fig. 6 in Ref. 57 and Fig. 11 in Ref. 7, our results exhibit the same kind of fundamental frequency dependence of pressure, where influence of subglottal pressure on fundamental frequency disappeared at higher pressures and small influence appeared at lower pressures. In accordance with demonstration and conclusion in prior studies,[7,57] causes of this phenomenon can be explained that the frequency of the linear oscillator depends only on the mass, the spring constant, and the damping constant, although the amplitude of the oscillation depends on driving force. The linear oscillator does not allow any variation in the frequency with the strength of the driving force, which is

directly proportional to the subglottal pressure.[57] To incorporate a variation in fundamental frequency with subglottal pressure into their model, Ishizaka *et al.*[7] introduced a nonlinear spring stiffness part depending on position with nonlinear coefficient $\eta$,[57] which is described in Eq. (A1). However, in this paper $\eta$ is assumed as zero (that is linear springs) in order to neglect the influence of nonlinear spring force. This assumption might cause that the spring stiffness or fundamental frequency is independent on variation in subglottal pressure, even if the amplitude of oscillation is high. In general, this result implicitly supports the suggestion from Ishizaka *et al.*[7] that "the fundamental frequency variations with different subglottal pressure might be ascribed to two causes: One is the collision of the vocal folds at closure when the amplitude of vibration is not too large and the subglottal pressure is small. Another is the nonlinearity of the deflection of the muscles and ligaments at large amplitudes of vibration and at subglottal pressures more than several cm $H_2O$." In other words, the small-amplitude vocal fold oscillation at small subglottal pressure and the nonlinearity/linearity (i.e., $\eta = 0$) of spring stiffness at higher subglottal pressure could cause different relationships between fundamental frequency and subglottal pressure.

Compared to the 3DM, 2D finite-element method (FEM) models[38,80] might indeed provide a more in detail representation concerning the physiological completeness.[12] Additionally, the detailed informations about the flow running through the human larynx can also be accurately simulated by FEM.[37] Nevertheless, the computation time for 1 cycle of vocal fold vibration using a fluid flow mechanical model presented in Ref. 81 takes about 15 h on one processor [AMD Opteron (TM), 1.8 MHz]. In studies on fluid-solid-acoustics interactions for human phonation,[82] the structure mechanical model of vocal folds takes about 20 days for 36 vibration cycles on four workstations (Intel Xeon 5160). The computational time for the 3DM takes currently 703–749 ms for 100 ms on a standard PC [Intel(R) Core(TM)2 Duo E8500, 3.16 GHz, 3.50 Gbyte RAM]. In comparison to FEM models, the 3DM purposefully has significantly less computational complexity, resulting in significantly higher computational speeds. As shown, by using the 3DM the mechanical properties not only from anterior to posterior but also in the vertical and lateral directions are simulated. Hence, it seems appropriate for extensive parameter studies and optimization on excised full/hemilarynx dynamics in future as performed with two or multi-mass-models so far.

## V. SUMMARY

To enhance our understanding of both normal and disordered voice production, a more complete description of vocal fold dynamics may be necessary than is offered by the two-mass-model of Ishizaka and Flanagan (1972). Recent research has revealed the 3D dynamics of the medial surface of the vocal folds across a number of laboratory models, including the excised human larynx, the *in vivo* canine larynx, and physical models. The 3DM was developed to help conceptualize these data into one coherent model of vocal fold vibration. Preliminary data from the 3DM demonstrate its ability to oscillate in patterns which are similar to available experimental data. In future experiments, it is hoped that tissue parameters of experimental data (including tissue masses and tissue stiffnesses along the entire medial surface of the vocal folds) may be deduced as the parameters of the 3DM are optimized to fit its dynamics to dynamical experimental data. In this way, a firmer connection may be established between biomechanical tissue properties and the resultant vocal fold dynamics.

## APPENDIX: DESCRIPTION OF APPLIED FORCES IN THE 3D MODEL

In following descriptions of the applied forces, the stiffness is defined as

$$k(x) = k(1 + \eta x^2),\qquad (A1)$$

where $k$ is assumed to be a linear constant. $\eta$ is a nonlinear coefficient and can be set to 1 mm$^{-2}$ for incorporating a nonlinear spring force into the model.[7,57] In this paper $\eta$ is set to 0 mm$^{-2}$, which indicates that the spring stiffness is independent from displacements.

### 1. Anchor force

The anchor force is defined as

$$\vec{\mathbf{F}}_{i,s}^A := -k_{i,s}^a \cdot (\|\mathbf{a}_{i,s} - \mathbf{x}_{i,s}^r\|_2 - \|\mathbf{a}_{i,s} - \mathbf{x}_{i,s}\|_2) \cdot \vec{\mathbf{u}}_{i,s}^a$$
$$- r_{i,s}^a \cdot (\vec{\mathbf{u}}_{i,s}^a)^T \cdot \dot{\mathbf{x}}_{i,s} \cdot \vec{\mathbf{u}}_{i,s}^a,$$

$$\vec{\mathbf{u}}_{i,s}^a = \frac{\mathbf{a}_{i,s} - \mathbf{x}_{i,s}}{\|\mathbf{a}_{i,s} - \mathbf{x}_{i,s}\|_2},\qquad (A2)$$

with $\mathbf{a}_{i,s}$ the 3D position of anchor; $\vec{\mathbf{u}}_{i,s}^a$ the unit vector (i.e., normalized vector) pointing from mass element position to adjacent anchor position in 3D, see Fig. 12(a); $k_{i,s}^a$ the stiffness of the vocal fold tissue; $r_{i,s}^a$ the anchor spring damping coefficient incorporating the viscous loss of the vocal fold tissue; with the damping ratio $\zeta_s$, the damping coefficient $r_{i,s}^a$ is expressed as[6,7,83]

$$r_{i,s}^a = 2\zeta_s \sqrt{m_{i,s} k_{i,s}^a};\qquad (A3)$$

$\mathbf{x}_{i,s}^r$ the corresponding rest position of mass element $m_{i,s}$. Here, $(\cdot)^T$ denotes the transposed vector.

### 2. Vertical coupling force

Within the 3DM, due to the five planes arranged in vertical direction, we assume the vertical internal forces being represented by aid of vertical springs. According to the com-

J. Acoust. Soc. Am., Vol. 127, No. 2, February 2010

Yang *et al.*: Modeling of three-dimensional vocal fold dynamics    1027

putation of vertical spring force within 2DM proposed by Schwarz et al.,[6] the vertical spring force in the 3DM is derived after

$$
\begin{aligned}
\vec{\mathbf{F}}_{i,s}^V := & -k_{i,s-1}^v \cdot (\|\mathbf{x}_{i,s-1}^r - \mathbf{x}_{i,s}^r\|_2 - \|\mathbf{x}_{i,s-1} - \mathbf{x}_{i,s}\|_2) \cdot \vec{\mathbf{u}}_{i,s-1}^v \\
& - r_{i,s-1}^v \cdot (\vec{\mathbf{u}}_{i,s-1}^v)^T \cdot (\dot{\mathbf{x}}_{i,s} - \dot{\mathbf{x}}_{i,s-1}) \cdot \vec{\mathbf{u}}_{i,s-1}^v \\
& - k_{i,s}^v \cdot (\|\mathbf{x}_{i,s+1}^r - \mathbf{x}_{i,s}^r\|_2 - \|\mathbf{x}_{i,s+1} - \mathbf{x}_{i,s}\|_2) \cdot \vec{\mathbf{u}}_{i,s+1}^v \\
& - r_{i,s}^v \cdot (\vec{\mathbf{u}}_{i,s+1}^v)^T \cdot (\dot{\mathbf{x}}_{i,s} - \dot{\mathbf{x}}_{i,s+1}) \cdot \vec{\mathbf{u}}_{i,s+1}^v,
\end{aligned}
$$

$$
\vec{\mathbf{u}}_{i,s-1}^v = \frac{\mathbf{x}_{i,s-1} - \mathbf{x}_{i,s}}{\|\mathbf{x}_{i,s-1} - \mathbf{x}_{i,s}\|_2}, \quad \vec{\mathbf{u}}_{i,s+1}^v = \frac{\mathbf{x}_{i,s+1} - \mathbf{x}_{i,s}}{\|\mathbf{x}_{i,s+1} - \mathbf{x}_{i,s}\|_2}, \quad \text{(A4)}
$$

where $k_{i,s-1}^v$, $k_{i,s}^v$ and $r_{i,s-1}^v$, $r_{i,s}^v$ are the stiffness and damping coefficients of the vertical coupling springs connected with mass element $m_{i,s}$. The directions of unit vectors (i.e., normalized vector) $\vec{\mathbf{u}}_{i,s-1}^v$ and $\vec{\mathbf{u}}_{i,s+1}^v$ are, respectively, from position of mass element $m_{i,s}$ to position of mass element $m_{i,s-1}$ and $m_{i,s+1}$, see Fig. 12(a).

### 3. Longitudinal coupling force

With the stiffness $k_{i-1,s}^l$, $k_{i,s}^l$ and the damping coefficients $r_{i-1,s}^v$, $r_{i,s}^v$ of the longitudinal coupling springs connected to mass element $m_{i,s}$ the corresponding longitudinal spring force $\vec{\mathbf{F}}_{i,s}^L$ can be defined in accordance to Schwarz et al.[6] as

$$
\begin{aligned}
\vec{\mathbf{F}}_{i,s}^L := & -k_{i-1,s}^l \cdot (\|\mathbf{x}_{i-1,s}^r - \mathbf{x}_{i,s}^r\|_2 - \|\mathbf{x}_{i-1,s} - \mathbf{x}_{i,s}\|_2) \cdot \vec{\mathbf{u}}_{i-1,s}^l \\
& - r_{i-1,s}^l \cdot (\vec{\mathbf{u}}_{i-1,s}^l)^T \cdot (\dot{\mathbf{x}}_{i,s} - \dot{\mathbf{x}}_{i-1,s}) \cdot \dot{\mathbf{u}}_{i-1,s}^l \\
& - k_{i,s}^l \cdot (\|\mathbf{x}_{i+1,s}^r - \mathbf{x}_{i,s}^r\|_2 - \|\mathbf{x}_{i+1,s} - \mathbf{x}_{i,s}\|_2) \cdot \vec{\mathbf{u}}_{i+1,s}^l \\
& - r_{i,s}^l \cdot (\vec{\mathbf{u}}_{i+1,s}^l)^T (\dot{\mathbf{x}}_{i,s} - \dot{\mathbf{x}}_{i+1,s}) \cdot \vec{\mathbf{u}}_{i+1,s}^l,
\end{aligned}
$$

$$
\vec{\mathbf{u}}_{i-1,s}^l = \frac{\mathbf{x}_{i-1,s} - \mathbf{x}_{i,s}}{\|\mathbf{x}_{i-1,s} - \mathbf{x}_{i,s}\|_2}, \quad \vec{\mathbf{u}}_{i+1,s}^l = \frac{\mathbf{x}_{i+1,s} - \mathbf{x}_{i,s}}{\|\mathbf{x}_{i+1,s} - \mathbf{x}_{i,s}\|_2}. \quad \text{(A5)}
$$

According to the assumption proposed by Wong[52] and Schwarz,[6] the stiffness $k_{i,s}^l$ and damping coefficient $r_{i,s}^l$ can be expressed in terms of stiffness and damping from the anchor spring

$$
k_{i,s}^l = \xi_k \cdot (k_{i,s}^a + k_{i+1,s}^a) \quad \text{with } \xi_k = 0.2, \quad \text{(A6)}
$$

$$
r_{i,s}^l = \xi_r \cdot (r_{i,s}^a - r_{i+1,s}^a) \quad \text{with } \xi_r = 0.06. \quad \text{(A7)}
$$

The directions of unit vectors (i.e. normalized vectors) $\vec{\mathbf{u}}_{i-1,s}^l$ and $\vec{\mathbf{u}}_{i+1,s}^l$ are from mass element $m_{i,s}$ toward mass element $m_{i-1,s}$ and $m_{i+1,s}$ in 3D, respectively, illustrated in Fig. 12(b).

### 4. Collision impact force

Based on the definition of collision proposed by Gunter[84] as displayed in Fig. 12(c), the collision force $\vec{\mathbf{F}}_{i,s}^C$ acting on mass $m_{i,s}$ in the 3DM is defined as[6]

$$
\vec{\mathbf{F}}_{i,s}^C := -\sum_{\forall i'} k_{i',i,s}^c d_{i',i,s}^p \vec{\mathbf{u}}_{i,s}^c, \quad \text{(A8)}
$$

$$
k_{i',i,s}^c := 3 \cdot \left( \frac{l_{i'+1,i,s}}{l_{i',i,s} + l_{i'+1,i,s}} k_{i',s}^a + \frac{l_{i',i,s}}{l_{i',i,s} + l_{i'+1,i,s}} k_{i'+1,s}^A \right), \quad \text{(A9)}
$$

where $\vec{\mathbf{u}}_{i,s}^c$ is the unit vector for collision impact force in 3D, its corresponding direction is from the position of mass element $m_{i,s}$ to the position of the adjacent anchor, see Fig. 12(c). The stiffness $k_{i',i,s}^c$ is expressed according to the corresponding derivation in the 2DM.[6] $i'$ denotes the index of mass element opposite $m_{i,s}$.

### 5. Driving force

The aerodynamic force, which causes the mass elements on both sides of the 3DM to vibrate, is called driving force.[6] It is produced by the glottal flow originating from the lung and complies with the Bernoulli law.[1,48,49,85,86] However, the driving force $\vec{\mathbf{F}}_{i,s}^D$ depends on not only the subglottal pressure $P^{\text{sub}}$ but also the geometric dimensions, i.e., thickness $h_s$ and vocal fold length $l_g$ as well as the rest position $\mathbf{x}_{i,s}^r$.

In order to better simulate and compute the Bernoulli forces acting on the vocal folds from inferior to superior through the whole larynx, besides the simplified-method based on the derivation from the 2DM,[6,8] a so-called detailed-method is also used in the 3DM tentatively. Moreover, through the comparison between both methods the validity and corresponding accuracy of the computation for the Bernoulli forces are verified.

#### a. The simplified-method

In the 3DM, the simplified-method for computation of the Bernoulli driving force is presented as an extension from the 2DM. The main difference of the driving force between the 3DM and the 2DM is the Bernoulli forces that act on all five planes in 3D. Therefore, the driving pressure $P_s$ at each plane is taken into account. Meanwhile, we assume that the subglottal pressure $P^{\text{sub}}$ and the supraglottal pressure $P^{\text{supra}}$ are constant, where $P^{\text{supra}} = 0$ in accordance with the assumption of a jet function for no supraglottal loading.[8,69,85]

Figures 13(a) and 13(b) illustrate the convergent and divergent glottal shapes, respectively. Due to the function of the jet formed by airstream detaching from the vocal fold surface, the pressure inside the glottis is higher than zero for a convergent glottis and less than zero (or close to zero, between the flow separation point and the glottis exit) for a divergent glottis.[48,69] Therefore, in accordance with 2DM (Ref. 6) and Steinecke et al.,[8] the pressure $P_s$ located at plane $s$ within the 3DM can be defined as

$$
\begin{aligned}
& P_s \\
& = \begin{cases} P^{\text{sub}} \cdot \left( 1 - \left( \frac{\max(0, a_{\min})}{a_s} \right)^2 \right) & \text{for } s < \arg\min_{s=1,\ldots,5}(a_s), \\ 0 & \text{for } s \geq \arg\min_{s=1,\ldots,5}(a_s), \end{cases}
\end{aligned}
\quad \text{(A10)}
$$

Yang et al.: Modeling of three-dimensional vocal fold dynamics

$$a_{\min} = \min(a_1, a_2, a_3, a_4, a_5), \tag{A11}$$

where $a_s$ is the glottal area with $s = 1, \ldots, 5$.

The driving force acting on every mass element can be resolved into two components: horizontal component $\vec{\mathbf{F}}_{i,s}^{D,\text{hor}}$ and vertical component $\vec{\mathbf{F}}_{i,s}^{D,\text{ver}}$. Following this analysis, the driving force within the 3DM is defined as

$$\vec{\mathbf{F}}_{i,s}^{D} = [\vec{\mathbf{F}}_{i,s}^{D,\text{hor}}, \quad \vec{\mathbf{F}}_{i,s}^{D,\text{ver}}]^T, \tag{A12}$$

$$\vec{\mathbf{F}}_{i,s}^{D,\text{hor}} = P_s \cdot w_{i,s} \cdot h_s \cdot \vec{\mathbf{u}}_{i,s}^d, \tag{A13}$$

$$\vec{\mathbf{F}}_{i,s}^{D,\text{ver}} = \begin{cases} P^{\text{sub}} \cdot w_{i,s} \cdot \sqrt{(x_{i,s}^r - x_{i,s})^2 + (y_{i,s}^r - y_{i,s})^2} \cdot \vec{\mathbf{u}}_z & \text{if } s = 1, \quad |x_{i,s}^r| < |x_{i,s}|, \\ 0 & \text{if } s = 1, \quad |x_{i,s}^r| \geqslant |x_{i,s}|, \\ P_{s-1} \cdot w_{i,s} \cdot \sqrt{(x_{i,s-1} - x_{i,s})^2 + (y_{i,s-1} - y_{i,s})^2} \cdot \vec{\mathbf{u}}_z & \text{if } s \neq 1, \quad |x_{i,s-1}| > |x_{i,s}|, \\ 0 & \text{if } s \neq 1, \quad |x_{i,s-1}| \leqslant |x_{i,s}|, \end{cases} \tag{A14}$$

where $x_{i,s-1}$, $y_{i,s-1}$ and $x_{i,s}$, $y_{i,s}$ as well as $x_{i,s}^r$, $y_{i,s}^r$ are the $x$- and $y$-components of $\mathbf{x}_{i,s-1}$, $\mathbf{x}_{i,s}$, $\mathbf{x}_{i,s}^r$. Additionally, $h_s$ indicates the thickness of mass element $m_{i,s}$, and $\vec{\mathbf{u}}_{i,s}^d$ is the unit vector for the horizontal component of driving force. The direction of the unit vectors is perpendicular to the vertical projection of $w_{i,s}$ at the horizontal plane. They point from inside the glottis towards outside the glottis, see Fig. 12(d). $\vec{\mathbf{u}}_z$ is the unit vector along the vertical axis.

In order to reasonably distribute the driving pressure acting on individual mass elements within the 3DM, an effective side length $w_{i,s}$ is used.[6] During phonation the position of collision point will be taken into account for the calculation of $w_{i,s}$, if the adjacent collision point occurs between mass element $m_{i,s}$ and mass element $m_{i-1,s}$ or $m_{i+1,s}$. Otherwise, $w_{i,s}$ is defined from the middle point of the adjacent longitudinal springs, i.e., $w_{i,s} = 0.5 \cdot \|\mathbf{x}_{i-1,s} - \mathbf{x}_{i+1,s}\|_2$, see Fig. 12(d).

### b. The detailed-method

The detailed-method is based on an approximate pressure distribution along the glottal flow.[7,66,87] Owing to the small dimensions of the human glottis and the high velocity of glottal flow, we can assume the glottal flow to be quasi-steady.[8] By using the Bernoulli equation for one-dimensional flow, the driving pressure distribution along the glottal flow can be obtained,[7] which is schematically illustrated in Fig. 13(c).

Pressure $P_{1,a}$ locates at the inferior glottal entrance corresponding to the lower edge of mass element $m_{i,1}$. According to the Bernoulli equation,[7,8,18] the approximate pressure distribution is described by

$$P^{\text{sub}} - P_{i,a} = \frac{\rho}{2} \cdot \left(\frac{U_g}{a_1}\right)^2, \tag{A15}$$

$$P_{s,a} - P_{s,b} = 12 \cdot \frac{\mu l_g^2 h_s}{a_s^3} \cdot U_g + \frac{\rho h_s}{a_s} \cdot \frac{dU_g}{dt}, \tag{A16}$$

$$P_{s+1,a} - P_{s,b} = \frac{\rho}{2} \cdot U_g^2 \cdot \left(\frac{1}{a_s^2} - \frac{1}{a_{s+1}^2}\right), \tag{A17}$$

where $P_{s,a}$, $P_{s,b}$ are the pressures at the lower and upper edges of mass element $m_{i,s}$ located at plane $s$. Additionally, $\rho$ and $\mu$ are the density and viscosity coefficients, of air, respectively. Due to the simplification of multi-mass-model as well as the small value of the air viscous losses (approximately $1.87 \times 10^{-7}$ g/cm ms), the air viscosity coefficient $\mu$ is predefined to zero in Eq. (A16). Hence, the consistency within the model development (in accordance with prior studies)[1,6,8,12] can be retained. $l_g$ is the vocal fold length, and $a_s$ is the cross-sectional glottal area at different planes ($s = 1, \ldots, 5$). Moreover, by using the above-proposed equations (A15)–(A17), the pressures $P_{s,a}$, $P_{s,b}$ at each plane can be derived. The volume flow rate of glottal airflow is $U_g$, which is derived in accordance with the assumption in Ref. 8 as follows:

$$U_g = \sqrt{\frac{2P^{\text{sub}}}{\rho}} a_{\min}. \tag{A18}$$

The mean value of the pressure acting on the mass element located at plane $s$ is given by

$$P_{s,m} = \tfrac{1}{2}(P_{s,a} + P_{s,b}). \tag{A19}$$

Bernoulli's law is less applicable to the divergent glottis in which the glottal pressure is less than zero.[48] In order to ensure the feasibility of Bernoulli's law pressure $P_{s,m}$ is set to zero, if $P_{s,m}$ derived in Eq. (A19) is negative.

Therefore, the horizontal component $\vec{\mathbf{F}}_{i,s}^{D,\text{hor}}$ and vertical component $\vec{\mathbf{F}}_{i,s}^{D,\text{ver}}$ of the driving force equation (A12) can be obtained in accordance to Lucero[18] and Koizumi[66] as

$$\vec{\mathbf{F}}_{i,s}^{D,\text{hor}} = P_{s,m} \cdot w_{i,s} \cdot h_s \cdot \vec{\mathbf{u}}_{i,s}^d, \tag{A20}$$

J. Acoust. Soc. Am., Vol. 127, No. 2, February 2010

Yang et al.: Modeling of three-dimensional vocal fold dynamics 1029

$$\vec{\mathbf{F}}_{i,s}^{D,\text{ver}} = \begin{cases} P^{\text{sub}} \cdot w_{i,s} \cdot \sqrt{(x_{i,s}^{r} - x_{i,s})^2 + (y_{i,s}^{r} - y_{i,s})^2} \cdot \vec{\mathbf{u}}_z & \text{if } s = 1 \\ P_{s-1,b} \cdot w_{i,s} \cdot \sqrt{(x_{i,s-1} - x_{i,s})^2 + (y_{i,s-1} - y_{i,s})^2} \cdot \vec{\mathbf{u}}_z & \text{if } s \neq 1, \quad s < \arg\min\limits_{s=1,\dots,5}(a_s) \\ 0 & \text{if } s \neq 1, \quad s \geq \arg\min\limits_{s=1,\dots,5}(a_s), \end{cases} \tag{A21}$$

where $P_{s-1,b}$ indicates the pressure at the upper edge of mass element $m_{i,s-1}$ at plane $s-1$, and $a_s$ denotes the cross sectional glottal area at different planes ($s=1,\dots,5$).

In order to simulate the 3D dynamics of human vocal folds, according to Eq. (1), the differential system of the new proposed model is solved by using the classical Runge–Kutta algorithm with the increment of 0.25 ms.[1,2]

[1]M. 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**, 773–781 (2002).

[2]R. 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**, 1099–1108 (2006).

[3]U. Eysholdt, F. Rosanowski, and U. Hoppe, "Irregular vocal fold vibrations caused by different types of laryngeal asymmetry," Eur. Arch. Otorhinolaryngol. **260**, 412–417 (2003).

[4]L. Sulica and A. Blitzer, "Vocal fold paresis: Evidence and controversies," Curr. Opin. Otolaryngol. Head Neck Surg. **15**, 159–162 (2007).

[5]N. Isshiki, M. Tanabe, K. Ishizaka, and D. Broad, "Clinical significance of asymmetrical vocal cord tension," Ann. Otol. Rhinol. Laryngol. **86**, 58–66 (1977).

[6]R. Schwarz, M. Döllinger, T. Wurzbacher, U. Eysholdt, and J. Lohscheller, "Spatiotemporal quantification of vocal fold vibrations using high-speed videoendoscopy and a biomechanical model," J. Acoust. Soc. Am. **123**, 2717–2732 (2008).

[7]K. Ishizaka and J. L. Flanagan, "Synthesis of voiced sounds from a two-mass model of the vocal cords," Bell Syst. Tech. J. **51**, 1233–1268 (1972).

[8]I. Steinecke and H. Herzel, "Bifurcations in an asymmetric vocal fold model," J. Acoust. Soc. Am. **97**, 1571–1578 (1995).

[9]P. Mergell, H. Herzel, and I. Titze, "Irregular vocal-fold vibration high-speed observation and modeling," J. Acoust. Soc. Am. **108**, 2996–3002 (2000).

[10]M. Döllinger, T. Braunschweig, J. Lohscheller, U. Eysholdt, and U. Hoppe, "Normal voice production: Computation of driving parameters from endoscopic digital high speed images," Methods Inf. Med. **42**, 271–276 (2003).

[11]T. 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**, 1012–1027 (2006).

[12]T. Wurzbacher, M. Döllinger, R. Schwarz, U. Hoppe, U. Eysholdt, and J. Lohscheller, "Spatiotemporal classification of vocal fold dynamics by a multi mass model comprising time-dependent parameters," J. Acoust. Soc. Am. **123**, 2324–2334 (2008).

[13]S. Adachi and J. Yu, "Two-dimensional model of vocal fold vibration for sound synthesis of voice and soprano singing," J. Acoust. Soc. Am. **117**, 3213–3224 (2005).

[14]Y. Zhang and J. J. Jiang, "Chaotic vibrations of a vocal fold model with a unilateral polyp," J. Acoust. Soc. Am. **115**, 1266–1269 (2004).

[15]I. R. Titze and B. H. Story, "Rules for controlling low-dimensional vocal fold models with muscle activation," J. Acoust. Soc. Am. **112**, 1064–1076 (2002).

[16]I. R. Titze, "Theoretical analysis of maximum flow declination rate versus maximum area declination rate in phonation," J. Speech Lang. Hear. Res. **49**, 439–447 (2006).

[17]J. D. Sitt, A. Amador, F. Goller, and G. B. Mindlin, "Dynamical origin of spectrally rich vocalizations in birdsong," Phys. Rev. E **78**, 011905 (2008).

[18]J. C. Lucero and L. L. Koenig, "Simulations of temporal patterns of oral airflow in men and women using a two-mass model of the vocal folds under dynamic control," J. Acoust. Soc. Am. **117**, 1362–1372 (2005).

[19]M. Döllinger, D. A. Berry, and D. W. Montequin, "The influence of epilarynx area on vocal fold dynamics," Otolaryngol.-Head Neck Surg. **135**, 724–729 (2006).

[20]I. R. Titze, "Parametrization of the glottal area, glottal flow, and vocal fold contact area," J. Acoust. Soc. Am. **75**, 570–580 (1984).

[21]M. Döllinger and D. A. Berry, "Visualization and quantification of the medial surface dynamics of an excised human vocal fold during phonation," J. Voice **20**, 401–413 (2006).

[22]I. R. Titze, "The physics of small-amplitude oscillation of the vocal folds," J. Acoust. Soc. Am. **83**, 1536–1552 (1988).

[23]I. R. Titze, S. S. Schmidt, and M. R. Titze, "Phonation threshold pressure in a physical model of the vocal fold mucosa," J. Acoust. Soc. Am. **97**, 3080–3084 (1995).

[24]R. W. Chan and I. R. Titze, "Dependence of phonation threshold pressure on vocal tract acoustics and vocal fold tissue mechanics," J. Acoust. Soc. Am. **119**, 2351–2362 (2006).

[25]D. A. Berry, M. J. O. Clark, D. W. Montequin, and I. R. Titze, "Characterization of the medial surface of the vocal folds," Ann. Otol. Rhinol. Laryngol. **110**, 470–477 (2001).

[26]A. Boessenecker, D. A. Berry, J. Lohscheller, U. Eysholdt, and M. Döllinger, "Mucosal wave properties of a human vocal fold," Acta. Acust. Acust. **93**, 815–823 (2007).

[27]M. Döllinger and D. A. Berry, "Computation of the three-dimensional medial surface dynamics of the vocal folds," J. Biomech. **39**, 369–374 (2006).

[28]M. Döllinger, N. Tayama, and D. A. Berry, "Empirical eigenfunctions and medial surface dynamics of a human vocal fold," Methods Inf. Med. **44**, 384–391 (2005).

[29]M. 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**, 3174–3183 (2005).

[30]M. 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**, 1646–1654 (2005).

[31]I. R. Titze, J. J. Jiang, and T. Y. Hsiao, "Measurement of mucosal wave propagation and vertical phase difference in vocal fold vibration," Ann. Otol. Rhinol. Laryngol. **102**, 58–63 (1993).

[32]D. 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**, 2539–2547 (2001).

[33]C. G. Garrett, J. R. Coleman, and L. Reinisch, "Comparative histology and vibration of the vocal folds: Implications for experimental studies in microlaryngeal surgery," Laryngoscope **110**, 814–824 (2000).

[34]J. Kobayashi, E. Yumoto, M. Hyodo, and K. Gyo, "Two-dimensional analysis of vocal fold vibration in unilaterally atrophied larynges," Laryngoscope **110**, 440–446 (2000).

[35]Z. Zhang, J. Neubauer, and D. A. Berry, "Aerodynamically and acoustically driven modes of vibration in a physical model of the vocal folds," J. Acoust. Soc. Am. **120**, 2841–2849 (2006).

[36]M. 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**, 1089–1103 (2008).

[37]S. Becker, S. Kniesburges, S. Müller, A. Delgado, G. Link, M. Kaltenbacher, and M. Döllinger, "Flow-structure-acoustic interaction in a human voice model," J. Acoust. Soc. Am. **125**, 1351–1361 (2009).

[38]F. Alipour-Haghighi, D. A. Berry, and I. R. Titze, "A finite-element model of vocal fold vibration," J. Acoust. Soc. Am. **108**, 3003–3012 (2000).

[39]D. A. Berry, H. Herzel, I. R. Titze, and K. Krischer, "Interpretation of biomechanical simulations of normal and chaotic vocal fold vibrations with empirical eigenfunctions," J. Acoust. Soc. Am. **95**, 3595–3604 (1994).

[40]A. 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).

[41]J. 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).

[42]O. 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).

[43]A. 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**, 175–189 (1996).

[44]P. 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**, 77–82 (2001).

[45]R. Schwarz, "Model-based quantification of pathological voice production," in *Kommunikationsstörungen, Berichte aus Phoniatrie und Pädaudiologie (Dysphonia, Reports from Phoniatrics and Pedaudiology)*, edited by U. Eysholdt (Shaker, Aachen, 2007).

[46]T. Wurzbacher, "Vocal fold dynamics-quantification and model-based classification," in *Kommunikationsstörungen, Berichte aus Phoniatrie und Pädaudiologie (Dysphonia, Reports from Phoniatrics and Pedaudiology)*, edited by U. Eysholdt (Shaker, Aachen, 2007).

[47]B. H. Story and I. R. Titze, "Voice simulation with a body-cover model of the vocal folds," J. Acoust. Soc. Am. **97**, 1249–1260 (1995).

[48]I. R. Titze, *Principles of Voice Production* (Prentice-Hall, Englewood Cliffs, NJ, 1994).

[49]J. L. Flanagan and K. Ishizaka, "Computer model to characterize the air volume displaced by the vibrating vocal cords," J. Acoust. Soc. Am. **63**, 1559–1565 (1978).

[50]M. C. Su, T. H. Yeh, C. T. Tan, C. D. Lin, O. C. Linne, and S. Y. Lee, "Measurement of adult vocal fold length," J. Laryngol. Otol. **116**, 447–449 (2002).

[51]H. Gray, in *Anatomy of the Human Body*, 20th ed., edited by W. H. Lewis (Lea & Febiger, Philadelphia, 1918).

[52]D. Wong, M. Ito, N. B. Cox, and I. R. Titze, "Observation of perturbations in a lumped-element model of the vocal folds with application to some pathological cases," J. Acoust. Soc. Am. **89**, 383–394 (1991).

[53]J. J. Jiang and I. R. Titze, "A methodological study of hemilaryngeal phonation," Laryngoscope **103**, 872–882 (1993).

[54]J. C. Lucero and L. L. Koenig, "On the relation between the phonation threshold lung pressure and the oscillation frequency of the vocal folds," J. Acoust. Soc. Am. **121**, 3280–3283 (2007).

[55]R. W. Chan, I. R. Titze, and M. R. Titze, "Further studies of phonation threshold pressure in a physical model of the vocal fold mucosa," J. Acoust. Soc. Am. **101**, 3722–3727 (1997).

[56]G. Farin, *Curves and Surfaces for Computer-Aided Geometric Design: A Practical Guide (Computer Science and Scientific Computing)* (Academic, San Francisco, CA, 1996).

[57]L. 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**, 386–393 (2006).

[58]U. Hoppe, F. Rosanowski, M. Döllinger, J. Lohscheller, and U. Eysholdt, "Visualization of the laryngeal motorics during a glissando," J. Voice **17**, 370–376 (2003).

[59]J. Lohscheller and U. Eysholdt, "Phonovibrogram visualization of entire vocal fold dynamics," Laryngoscope **118**, 753–758 (2008).

[60]Y. Yan, K. Ahmad, M. Kunduk, and D. Bless, "Analysis of vocal-fold vibrations from high-speed laryngeal images using a hilbert transform-based methodology," J. Voice **19**, 161–175 (2005).

[61]T. Braunschweig, J. Flaschka, P. Schelhorn-Neise, and M. Döllinger, "High-speed video analysis of the phonation onset, with an application to the diagnosis of functional dysphonias," Med. Eng. Phys. **30**, 59–66 (2008).

[62]T. Braunschweig, P. Schelhorn-Neise, and M. Döllinger, "Diagnosis of functional voice disorders by using the high speed recording technics," Laryngorhinootologie **87**, 323–330 (2008).

[63]D. D. Deliyski, P. P. Petrushev, H. S. Bonilha, T. T. Gerlach, B. Martin-Harris, and R. E. Hillman, "Clinical implementation of laryngeal high speed videoendoscopy: Challenges and evolution," Folia Phoniatr Logop **60**, 33–44 (2008).

[64]M. Döllinger, "Parameter estimation of vocal fold dynamics by inversion of a biomechanical model," in *Kommunikationsstörungen, Berichte aus Phoniatrie und Pädaudiologie, Band 10 (Dysphonia, Reports from Phoniatrics and Pedaudiology)*, edited by U. Eysholdt (Shaker, Aachen, 2002).

[65]S. Y. Lowell and B. H. Story, "Simulated effects of cricothyroid and thyroarytenoid muscle activation on adult-male vocal fold vibration," J. Acoust. Soc. Am. **120**, 386–397 (2006).

[66]T. Koizumi, S. Taniguchi, and S. Hiromitsu, "Two-mass models of the vocal cords for natural sounding voice synthesis," J. Acoust. Soc. Am. **82**, 1179–1192 (1987).

[67]X. 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**, 625–641 (2009).

[68]J. J. Jiang and I. R. Titze, "Measurement of vocal fold intraglottal pressure and impact stress," J. Voice **8**, 132–144 (1994).

[69]S. L. Thomson, L. Mongeau, and S. H. Frankel, "Aerodynamic transfer of energy to the vocal folds," J. Acoust. Soc. Am. **118**, 1689–1700 (2005).

[70]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**, 519–531 (2007).

[71]N. A. George, F. F. M. de Mul, Q. Qiu, G. Rakhorst, and H. K. Schutte, "Depth-kymography: High-speed calibrated 3d imaging of human vocal fold vibration dynamics," Phys. Med. Biol. **53**, 2667–2675 (2008).

[72]T. Baer, "Observations of vocal fold vibration: Measurement of excised larynges," in *Vocal Fold Physiology*, edited by K. N. Stevens and M. Hirano (University of Tokyo Press, Tokyo, 1981), Chap. 10, pp. 119–133.

[73]J. B. Park and L. Mongeau, "Experimental investigation of the influence of a posterior gap on glottal flow and sound," J. Acoust. Soc. Am. **124**, 1171–1179 (2008).

[74]J. B. Park and L. Mongeau, "Instantaneous orifice discharge coefficient of a physical, driven model of the human larynx," J. Acoust. Soc. Am. **121**, 442–455 (2007).

[75]J. S. Drechsel and S. L. Thomson, "Influence of supraglottal structures on the glottal jet exiting a two-layer synthetic, self-oscillating vocal fold model," J. Acoust. Soc. Am. **123**, 4434–4445 (2008).

[76]C. Tao, J. J. Jiang, and Y. Zhang, "Simulation of vocal fold impact pressures with a self-oscillating finite-element model," J. Acoust. Soc. Am. **119**, 3987–3994 (2006).

[77]I. R. Titze, "On the relation between subglottal pressure and fundamental frequency in phonation," J. Acoust. Soc. Am. **85**, 901–906 (1989).

[78]J. J. Jiang, A. G. Shah, M. M. Hess, K. Verdolini, F. M. Banzali, and D. G. Ganson, "Vocal fold impact stress analysis," J. Voice **15**, 4–14 (2001).

[79]I. R. Titze, "Phonation threshold pressure: A missing link in glottal aerodynamics," J. Acoust. Soc. Am. **91**, 2926–2935 (1992).

[80]H. E. Gunter, "A mechanical model of vocal-fold collision with high spatial and temporal resolution," J. Acoust. Soc. Am. **113**, 994–1000 (2003).

[81]H. Luo, R. Mittal, X. Zheng, S. A. Bielamowicz, R. J. Walsh, and J. K. Hahn, "An immersed-boundary method for flowstructure interaction in biological systems with application to phonation," J. Comput. Phys. **227**, 9303–9332 (2008).

[82]G. Link, M. Kaltenbacher, M. Breuer, and M. Döllinger, "A 2D finite-element scheme for fluid-solid-acoustic interactions and its application to human phonation," Comput. Methods Appl. Mech. Eng. **198**, 3321–3334 (2008).

[83]J. L. Flanagan and L. Landgraf, "Self-oscillating source for vocal-tract synthesizers," IEEE Trans. Audio Electroacoust. **16**, 57–64 (1968).

[84]H. E. Gunter, "Modeling mechanical stresses as a factor in the etiology of benign vocal fold lesions.," J. Biomech. **37**, 1119–1124 (2004).

[85]M. Little, P. McSharry, I. Moroz, and S. Roberts, "A simple, quasi-linear, discrete model of vocal fold dynamics," in *Nonlinear Analyses and Algorithms for Speech Processing*, edited by M. Faundez-Zanuy, L. Janer, A. Esposito, A. Satue-Villar, J. Roure, and V. Espinosa-Duro (Springer, Berlin, 2005), Vol. **3817**, pp. 348–356.

[86]M. P. de Vries, H. K. Schutte, A. E. P. Veldman, and G. J. Verkerke, "Glottal flow through a two-mass model: Comparison of Navier-Stokes solutions with simplified models," J. Acoust. Soc. Am. **111**, 1847–1853 (2002).

[87]T. Ikeda and Y. Matsuzaki, "A two-dimensional vocal cord model for heavy voice," Trans. Jpn. Soc. Mech. Eng., Ser. C **63**, 693–699 (1997).