Acoustic properties of symmetric and asymmetric vocal fold vibration ⊘

Special Collection: Flow and Phonation

Qilin Liu ; Guofeng He ; Lea Sayce ; Haoxiang Luo ; Zheng Li ✉

🔴 Check for updates

View Online   Export Citation

Articles You May Be Interested In

Computational hemodynamics and hemoacoustic study on carotid bifurcation: Effect of stenosis and branch angle

Physics of Fluids (April 2024)

Influences of serrated trailing edge on the aerodynamic and aeroacoustic performance of a flapping wing during hovering flight

Physics of Fluids (January 2022)

Acoustic properties of symmetric and asymmetric vocal fold vibration

View Online   Export Citation   CrossMark

Qilin Liu,[1,a)] (iD) Guofeng He,[1] (iD) Lea Sayce,[2] (iD) Haoxiang Luo,[3] (iD) and Zheng Li[1,b)] (iD)

AFFILIATIONS

[1]Department of Mechatronics Engineering, Morgan State University, Baltimore, Maryland 21251, USA

[2]Department of Communication Science and Disorders, School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA

[3]Department of Mechanical Engineering, Vanderbilt University, Nashville, Tennessee 37240, USA

Note: This paper is part of the Special Topic, Flow and Phonation.
a)Electronic mail: 1992134173@qq.com
b)Author to whom correspondence should be addressed: zheng.li@morgan.edu

ABSTRACT

To investigate the acoustic properties of signals generated by symmetric and asymmetric vocal fold vibrations, a flow-acoustic splitting method is employed to model the glottal airflow associated with voice production. The perturbed compressible pressure, p', is calculated by the linearized perturbed compressible equations (LPCE). Based on p' and the source term of the LPCE, acoustic behavior related to the medial thickness of the vocal fold and the frequency difference between the two sides of the vocal fold are analyzed. The results show that the opposite-polarity source pair is responsible for the production of p' and the opposite-polarity source pair is located right at the entrance of the glottal gap. The frequency difference of the two sides diminishes the opposite-polarity source pair and causes amplitude modulation of p'. Consequently, asymmetric vibration can lead to voice problems. The increase in the medial thickness makes the distribution of the paired sources more compact and stronger, and it enhances the intensity ratio between the p' and the hydrodynamic pressure variation, thereby positively contributing to voice production.

Published under an exclusive license by AIP Publishing. https://doi.org/10.1063/5.0279609

I. INTRODUCTION

Voice production is critical to the daily communication and social life of humans. The human voice results from a complex interplay of the vibration of the vocal fold and exhaled air pushed through the glottis, phonation, in addition to the shaping of sound in the vocal tract, i.e., resonance, or manipulation of acoustic structures within the vocal tract, i.e., articulation. Abnormal vocal fold motion can lead to voice disorders, which may result in negative professional, social, and emotional consequences (Nerrière *et al.,* 2009). The study of vocal fold motion and its acoustic properties is necessary to inform treatment outcomes and expectations for patients with dysphonia.

Abnormalities in vocal fold vibration can arise from various factors. One common cause is vocal fold paralysis, which can affect one (unilateral) or both (bilateral) vocal folds. In such cases, asymmetric vibration of the vocal fold is the most common clinical manifestation (Chung *et al.,* 2024). Surgical interventions aim to medialize the vocal fold by augmentation with implants (Zhang *et al.,* 2020), injectables (Henriques *et al.,* 2023), or designing laryngeal prostheses.

(Kosako *et al.,* 2023). Factors resulting from these interventions such as local inflammation (Enver *et al.,* 2021), fibrosis (Woo *et al.,* 2001), implant extrusion (Konjević *et al.,* 2024), etc., may influence treatment efficacy and undesirable outcomes in short- and long-terms.

Experiments to understand vocal fold vibration have been conducted using biological and synthetic models (Michaud-Dorko *et al.,* 2024). These experiments help identify the factors related to vibratory asymmetry and their impact on the voice. Li *et al.* (2021) performed *in vivo* phonation tests in a New Zealand white rabbit model. In their experiments, one side of the vocal fold is suture-adducted to model unilateral vocal fold paralysis (UVFP), and high-speed videoendoscopy (HSV) is utilized to capture vocal fold vibration in the UVFP conditions. Their results show that the two sides of the vocal fold can vibrate asymmetrically, with the adducted side vibrating at a 10% higher frequency than the resting side. In experiments using a synthetic vocal fold model, Pickup and Thomson (2009) investigated the impact of asymmetrical stiffness in vocal fold on the vibration pattern. Their results indicate that asymmetrical stiffness leads to differences in

vibrating frequency, which may impact voice production. A similar outcome has been observed in medialization laryngoplasty, a surgical procedure that involves inserting a silicone-elastomer implant into the vocal fold. As reported by Zhang et al. (2020), an increase in implant stiffness raises phonation frequency, which may become excessively high and undesirable. This suggests that the frequency difference between a pair of vocal folds is a crucial factor in voice outcomes.

In addition to vibrating frequency, the shape of the medial surface also plays a significant role in medialization laryngoplasty (Zhang, 2018; Zhang and Chhetri, 2019; and Zhang, 2024). The medial surface is the region where two sides of the vocal fold contact each other. The vertical length of medial surfaces can be quantified as the medial thickness. Zhang et al. (2020) investigated vocal fold deformation resulting from implant insertion. Their experiments show that a soft implant, with a Young's Modulus of 11 kPa, conformed to the shape of the vocal fold, while a stiff implant, with a Young's Modulus of 1386 kPa, maintained its original shape. Whether the implant is soft or stiff, both types led to a significant increase in medial thickness, rising from 2 mm to about 4 mm, when the implants occupied a large space previously filled by the original tissue. Zhang and Chhetri (2019) modified the medial thickness of excised human larynges using a wooden stick to analyze the acoustic effects. Their results indicate that increased medial thickness enhances vocal fold closure and produces higher-order harmonics in the voice. Therefore, both stiffness differences and medial thickness are crucial factors closely related to the vocal fold motion and voice outcomes.

Compared to in vivo and ex vivo experiments, numerical models offer more detailed insights into this complex process involving flow-structure-acoustic interactions (FSAI) (Falk et al., 2021; Maurerlehner et al., 2022; McCollum, et al., 2023a; 2023b; and Zhang, 2024). Current efforts to develop computational models of voice production have been reviewed by Döllinger et al. (2023). Due to the low Mach number (<0.3) characteristic, the FSAI model commonly employs the hydrodynamic/acoustic splitting method. This approach allows the FSAI model to be divided into a flow-structure interaction (FSI) model and an acoustic field model (He et al., 2022, 2024). For the FSI model, direct numerical simulation (DNS) and large-eddy simulation (LES) are two kinds of advanced computational techniques for the flow simulation. DNS needs to resolve the Kolmogorov length scale, the smallest scale in flow (Choi and Moin, 2012), which is computationally expensive. Thus, DNS is employed for simplified vocal fold model, in which the flow is generally laminar (Bae and Moon, 2008). LES models the effect of small-scale and resolves the large-scale flow feature. Thus, LES is usually employed to simulate the turbulent flow in the vocal fold (Schoder et al., 2020; Falk et al., 2021). By employing the FSI model, geometrical parameters of the vocal fold have been extensively investigated. McCollum et al. (2023b) analyzed laryngeal dimensions in relation to sex differences. Their study examined how the vocal fold length, depth, and thickness influence the vocal fold deformation, excluding medial thickness. Zhang (2017, 2018) investigated the effects of medial thickness and vocal fold stiffness. However, his model assumes a symmetrical vocal fold motion about the midline, which does not account for asymmetric vibrations. For the asymmetry vocal fold motion, Xue et al. (2010) developed a two-dimensional model representing a pair of vocal folds. The motion was simulated using a two-mass model with different spring constants for the two sides, while the airflow was treated as incompressible. This model was able to capture

asymmetries in both vibration and the flow field. Their results show that the fundamental frequency of phonation is governed by the properties of the softer side of the vocal fold. Thus, the FSI system appears a selection of preferential frequency. However, the effect of medial thickness during asymmetric motion was not considered in these models, despite experimental studies highlighting its significance in voice production.

The flow-acoustic splitting method is effective in handling the scale disparity between the acoustic and flow fields (Maurerlehner et al., 2022). The acoustic analogy method and the acoustic perturbation approach are two common splitting methods. The acoustic analogy method predicts the far-field acoustics based on equivalent sound sources derived from the flow field. Zhao et al. (2002) employed the acoustic analogy method to predict the sound generated by a simplified axisymmetric vocal fold. Their results showed that the dipole equivalent sound source dominates when the unsteady surface force on the vocal fold is considered, with the strength of the dipole source being two orders of magnitude higher than that of the quadrupole source. Kosako et al. (2023) estimated the sound of a voice prosthesis using Lighthill's tensor. Zhang (2017, 2018) and Bodaghi et al. (2021), used the pressure near the exit boundary of the flow simulation as a monopole source to estimate the sound signals far from the vocal fold. However, while the acoustic analogy provides a mathematical description of sound sources, it lacks a direct relationship between the equivalent sources and the flow features. Thus, the equivalent sources may not really exist (Tam et al., 2008), making it difficult to investigate sound production, particularly in the near-source region.

The acoustic perturbation approach divides the computation into a linear superposition of the incompressible fluid field and the compressible perturbed field. There are many variations in the acoustic perturbation approach family, such as acoustic perturbation equations variant 2 (APE-2) (Ewert and Schröder, 2003), perturbed convective wave equation (PCWE) (Kaltenbacher et al., 2016; Maurerlehner et al., 2022; and Schoder, 2025), and linearized perturbed compressible equations (LPCE) (Seo and Moon, 2006; Seo and Mittal, 2011). The PCWE reformulates the APE-2 into a one-scale wave equation, thus computationally efficient. However, the PCWE needs a mean flow from a given flow simulation, which causes difficulty in parallel computing with the flow simulation (Döllinger et al., 2023). It also poses challenges in coupling with the movement of the vocal fold (Falk et al., 2021; Döllinger et al., 2023). By introducing an acoustic potential function into the original PCWE, Schoder (2025) attempted to solve the PCWE based solely on the instantaneous flow field which is also suitable for modeling acoustics on deformed grids for the FSI problems. In comparison with APE-2 and original-PCWE, the LPCE can be solved in parallel with the flow simulation and allows coupling with the moving boundary (Bae and Moon, 2008, 2011). The LPCE is derived by subtracting the incompressible Navier–Stokes (INS) equation from the fully compressible Navier–Stokes equation and then linearizing the perturbation equation (Seo and Moon, 2006). The LPCE has been validated by canonical problems (Seo and Moon, 2006; Seo and Mittal, 2011). Bae and Moon (2008) employed the INS/LPCE method to model an axisymmetric vocal fold. In their model, the INS equations are solved to provide the source term for the LPCE, with the vocal fold motion prescribed by a time-dependent contour function. A moving coordinate system is used to couple the vocal fold motion with the LPCE. Both the vortical flow field and the perturbed compressible pressure are

25 August 2025 14:22:32

analyzed to explore the relationship between voice quality and the vortical structure. Bodaghi, et al. (2021) successfully employed the INS/LPCE method to examine supraglottic acoustics in a two-dimensional computational model. These results demonstrate the suitability of the INS/LPCE method for vocal fold simulations. It should be note that although the LPCE can be coupled with the LES, in which the filtered INS is solved (Hwang and Moon, 2017). It is theoretically appropriate to couple the LPCE with the INS. Therefore, the INS/LPCE method will be used in the present study.

Motivated by experimental observations of vocal fold vibrating frequency and medial surface (Pickup and Thomson, 2009; Zhang et al., 2020; and Zhang and Chhetri, 2019). In the present paper, by coupling flow modeling and LPCE, both the glottal airflow and acoustic fields are analyzed to examine the sensitivity of medial thickness and frequency differences on the voice production. This provides insights into the coupling mechanisms of flow pulsatility and the acoustic properties of the asymmetric vocal fold motion.

II. PHYSICAL MODELS

In the current study, we focus on the acoustic signal in related various vocal fold motions. Therefore, to reduce the computational cost, we consider a two-dimensional one-way coupling approach, where the vocal fold motion is specified based on previous observations (Zhao et al., 2002; Bae and Moon, 2008). Although this approach does not resolve the tissue behavior in the vocal fold, it allows the motion of the medial surface to be parameterized, enabling a systematic study of vocal fold motion. Otherwise, due to the highly nonlinear nature of tissue behavior, employing a fully coupled approach will be computationally prohibitive for such an investigation. The flow field and vocal fold motion from the one-way coupling simulation will be used to solve the LPCE under various vibration conditions.

Figure 1 shows the geometry of the vocal fold. In Fig. 1(a), a cross-sectional view of the trachea shows the air flow being obstructed by a pair of vocal folds. Readers who are interested in anatomy view are encouraged to refer our previous work Li (2023). The vocal fold can be asymmetrical to the centerline, thus a 2D geometry is employed to model the vocal fold, with a pair of vocal folds labeled as fold-A and fold-B, as shown in Fig. 1(b). The width of the trachea is marked
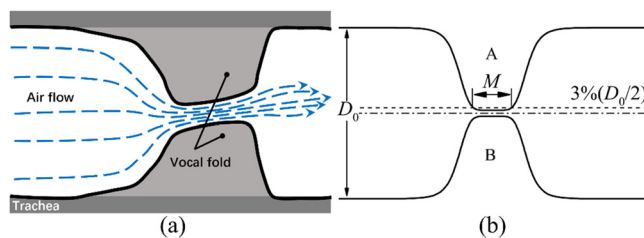
by D_0. The medial thickness, denoted by M, is the width of the vocal fold's medial surface. This kind of 2D geometry similar to that in present study has been widely used and could be found in other useful phonation models (Bodaghi et al., 2021; Bae and Moon, 2008; and Zhao et al., 2002). The entire computational domain is shown in Fig. 2. The airflow runs through the glottis from left to right. The region within range of $x/D_0 \in [-1, 1]$ is deformable to model the vocal fold motion. The downstream channel representing the trachea is long enough to encompass both the flow and acoustic fields. The computational domain consists of a physical domain and two sponge zones. The sponge zones are used to absorb sound waves, and their details will be further discussed in Sec. III B.

The airflow in the channel is driven by a pressure difference $\Delta p = 1$ kPa, which is in a typical range of human phonation (Holmberg and Perkell, 1989; Döllinger et al., 2023). The flow can be regarded as incompressible fluid for low Mach number (<0.3), as in the case of human phonation. The density of the airflow is set as $\rho_\infty = 1130$ g/cm^3 at body temperature and ambient pressure $p_\infty = 1.013 \times 10^5$ Pa. To reduce computational cost, the kinematic viscosity of the flow is set to $\nu_\infty = 2.04 \times 10^{-4}$ m^2/s, which is 10 times of the value in normal human phonation. Similar treatment has been widely used in other vocal fold models (Zhao et al., 2002; Bae and Moon, 2008; and McCollum, et al., 2023a; 2023b). Such treatment affects the turbulence flow and the high-frequency signals that are not the focus of this study. The model aims to investigate the primary acoustic behaviors resulting from vocal fold motion.

III. INS/LPCE MODEL AND VALIDATION
A. Governing equation

In the INS/LPCE method, the flow field is governed by the INS equation, which can be written as follows:

$$\begin{cases} \dfrac{\partial U_j}{\partial x_j} = 0 \\ \dfrac{\partial U_i}{\partial t} + \dfrac{\partial U_i U_j}{\partial x_j} = -\dfrac{1}{\rho_\infty}\dfrac{\partial P}{\partial x_i} + \dfrac{\partial}{\partial x_j}\left(\nu_\infty\left(\dfrac{\partial U_j}{\partial x_i} + \dfrac{\partial U_i}{\partial x_j}\right)\right), \end{cases} \quad (1)$$

where ρ_∞, U_i, P are incompressible quantities of density, velocity, and pressure, respectively. The INS and LPCE are forward coupled. After ρ_∞, U_i, P are obtained by solving the INS, they are substituted into the LPCE to estimate the perturbed compressible quantities of density, ρ', velocity, u'_i, and pressure, p'. p' contains radiative components that propagate as an acoustic wave, as discussed in the subsection 4 of Appendix. Thus, p' is employed to represent the acoustic signal in the present paper. Note that, to couple with the LPCE, P should be the absolute pressure. The incompressible pressure relative to the ambient pressure is defined by $P_r = P - p_\infty$. The LPCE is written as follows (Seo and Moon, 2006)

FIG. 1. Sketch of the vocal fold. (a) Air flow in an asymmetric vocal fold. (b) Two-dimensional vocal fold model.

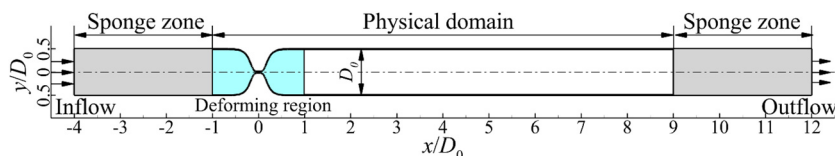**FIG. 2.** Computational domain of vocal fold model. The dash-dotted line marks the centerline.

$$
\begin{cases}
\dfrac{\partial \rho'}{\partial t} + U_j \dfrac{\partial \rho'}{\partial x_j} + \rho_\infty \dfrac{\partial u'_j}{\partial x_j} = 0 \\[2mm]
\dfrac{\partial u'_i}{\partial t} + \dfrac{\partial U_j u'_j}{\partial x_i} + \dfrac{1}{\rho_\infty} \dfrac{\partial p'}{\partial x_i} = 0 \\[2mm]
\dfrac{\partial p'}{\partial t} + U_j \dfrac{\partial p'}{\partial x_j} + \gamma P \dfrac{\partial u'_j}{\partial x_j} + u'_j \dfrac{\partial P}{\partial x_j} = -\dfrac{DP}{Dt},
\end{cases}
\tag{2}
$$

where $\gamma = 1.4$ is the ratio of specific heat, DP/Dt represents the material derivative of the pressure, i.e.,

$$
\frac{DP}{Dt} = \left(\frac{\partial P}{\partial t} + U_j \frac{\partial P}{\partial x_j} \right)
\tag{3}
$$

DP/Dt is the only source term in the LPCE, and it consists of a temporal component, $\partial P/\partial t$, and a convection component, $U_j(\partial P/\partial x_j)$. It In Eqs. (1) and (2), subscripts i and j are 1 or 2, representing the Cartesian components.

To describe the deformation of the vocal fold, a pair of contour functions is employed. The corresponding deforming region, as shown in Fig. 1, is depicted by a moving computational mesh. The vocal fold motion involves both medial-lateral oscillation and rotation of the medial surface, which is characteristic for the mucosal wave-like motion of the human vocal fold (Döllinger *et al.*, 2023). This kind of vocal fold motion similar to that in the present study has widely been used and could be found in other useful phonatory models (Šidlof *et al.*, 2015; Jo *et al.*, 2016; and Lasota *et al.*, 2023). A time-dependent contour function proposed by Zhao *et al.* (2002) is employed to model the vocal fold motion. To model the asymmetrical motion, the contour function is applied to each fold individually.

This function, denoted by $h(x, t)$, is written as follows (Bae and Moon, 2008; Zhao *et al.*, 2002)

$$
h(x, t) = \frac{D_0 + D_{\min}}{4} + \frac{D_0 - D_{\min}}{4} \tanh(s(x)) + \frac{1}{2}(1 - \tanh(s(x))) \\
\cdot (\beta_1 \cdot (x + x_c) - \beta_2 \cdot (x - x_c)),
\tag{4}
$$

where $D_{\min}$ and $D_{\max}$ are the minimum and maximum gaps of the glottis, respectively. $D_0 = 5D_{\max}$ is the width of the trachea, and $x_c = 0.42D_{\max}$ is a constant. The function $s(x)$ is defined as follows:

$$
s(x) = H \cdot |x|/D_{\max} - H \cdot D_{\max}/|x|,
\tag{5}
$$

where H is an adjustable coefficient. It relates to the curvature of the hyperbolic tangent function so it can be used to adjust the medial thickness. In Eq. (4), temporal coefficients $\beta_1 = \beta(t)$ and $\beta_2 = \beta(t + T/9)$ are given as follows (Bae and Moon, 2008)

$$
\beta(t) =
\begin{cases}
0, & t \le \dfrac{T}{9} \\[2mm]
0.244 \cdot \left(1 - \cos\left(\dfrac{9\pi}{4}\left(\dfrac{t}{T} - \dfrac{1}{9}\right)\right)\right), & \dfrac{T}{9} < t \le \dfrac{5T}{9} \\[2mm]
0.488, & \dfrac{5T}{9} < t \le \dfrac{6T}{9} \\[2mm]
0.244 \cdot \left(1 - \cos\left(3\pi\left(\dfrac{t}{T} - \dfrac{6}{9}\right)\right)\right), & \dfrac{6T}{9} < t \le T,
\end{cases}
\tag{6}
$$

where T is the period of vocal fold vibration. The vibration frequency is thus $f = 1/T$. A visualization of the vocal fold motion is shown in
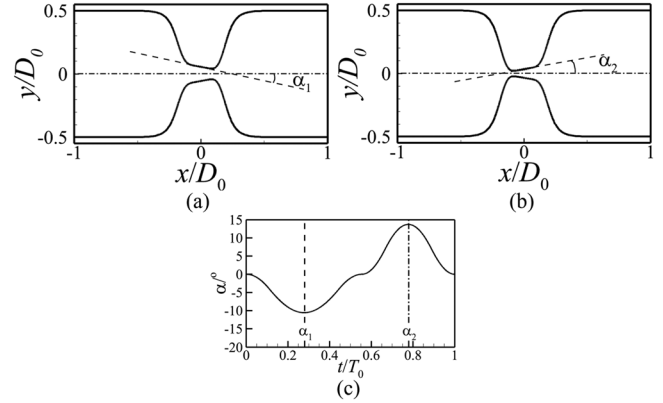

FIG. 3. Contours of the symmetric vocal fold with $D_{\min}/D_0 = 0.036$, $H = 2.8$. (a) Convergent contour at angle $\alpha = \alpha_1$. (b) Divergent contour at $\alpha = \alpha_2$. (c) Variation of α in a cycle of motion.

Fig. 3 by contour lines. The tilt angle of medial surfaces is estimated by $\alpha = \tan^{-1}(\beta_1 - \beta_2)$ and shown in Fig. 3(c) for a cycle of the vocal fold motion. The medial surface exhibits a convergent contour in Fig. 3(a) and a divergent contour in Fig. 3(b), mimicking the physiological motion of the medial surface. The mesh displacement in the deformable region (Fig. 2) is done by moving the mesh points in the lateral direction and in a linear fashion. The streamwise displacement is set to zero. That is,

$$
\begin{cases}
\Delta x = 0 \\[2mm]
\Delta y = y_0 \cdot \dfrac{\Delta h}{h(x_0, t_0)},
\end{cases}
\tag{7}
$$

where (x_0, y_0) is the coordinate of point P_0 in the deformable region at initial time t_0, $\Delta h = h(x_0, t) - h(x_0, t_0)$ is the displacement of the contour line at $x = x_0$ and time t, Δx and Δy are the streamwise and lateral displacements of point P_0 at time t.

B. Boundary conditions and numerical treatment

For the vocal fold model, the outlet pressure is fixed at the ambient pressure p_∞, the inlet pressure is 1 kPa higher than the outlet pressure to establish a driving pressure. A zero normal velocity gradient condition is applied for both the inlet and outlet boundaries. At the channel wall, a moving-wall condition is applied, where the velocity at the wall and the normal pressure gradient are given as follows:

$$
U_i = \frac{\partial x_{w,i}}{\partial t}, \quad \frac{\partial P}{\partial x_i} \cdot n_i = 0,
\tag{8}
$$

where $x_{w,i}$ and n_i are the Cartesian components of the position vector and the wall normal vector, respectively.

For the LPCE, the wall boundary conditions are (Bae and Moon, 2011)

$$
u'_i \cdot n_i = 0, \quad \frac{\partial p'}{\partial x_i} \cdot n_i = 0, \quad \frac{\partial \rho'}{\partial x_i} \cdot n_i = 0.
\tag{9}
$$

These conditions allow pressure waves to reflect at the wall, corresponding to an acoustically hard boundary (Kaltenbacher, 2018).

25 August 2025 14:22:32

When solving the LPCE, numerical instability may generate grid-dependent oscillations, and numerical reflections may happen at the inlet and outlet boundaries (Freund, 1997; Xu and Yang, 2021). These nonphysical waves can ruin the solution and must be eliminated. In the present model, a damping term and a convection term (Freund, 1997) are added on the left side of Eq. (2) to minimize the nonphysical reflection waves near the inlet/outlet. In addition to the boundary treatment, artificial viscosity is employed for the LPCE to dampen the wiggles in the computational domain. Details about these numerical treatments are provided in the subsection 1 of Appendix.

C. Numerical method

The governing equations, boundary conditions, and numerical treatments are implemented in the commercial software COMSOL Multiphysics to conduct the simulation. In the software, Eq. (1) is solved using the laminar flow module to obtain the incompressible flow field. Equation (2), is customized in the software and solved using the user-defined PDE module. The displacement field, as defined in Eq. (7), is coded into the software by the moving mesh module. For spatial discretization, quadratic elements are used to discretize the domain. For time marching, the generalized-α method is employed which is an implicit method and can properly deal with the large differences in the speed between the flow and the acoustic wave. At each time step, Eqs. (1) and (2) are solved in sequence by the constant Newton iteration method, with the PARDISO solver applied for each iteration. The convergence criteria of the iteration is set to 10^{-6}. In such a way, the INS and LPCE equations are directly solved using the finite element method.

D. Model validation

To validate the model capability in simulating the incompressible flow induced sound problem, a straight duct with a pulsating velocity inlet is simulated. This test case mimics the trachea and has been used to validate the INS/LPCE solver (Bodaghi, et al., 2021). Computational domain and mesh of the duct are shown in Fig. 4. The duct is in size of 2.8×17.3 cm, which is meshed by 20×70 grid cells. Perturbations outside of the duct propagate freely in a rectangular domain which is in size of 52.5×35 cm and meshed by 160×70 grid cells. In Fig. 3, a sponge zone is applied to absorb nonphysical reflection waves near the boundary. A Gaussian velocity pulse is prescribed at the inlet boundary, defined as

$$U_2 = 6[\text{cm/s}] \cdot \exp\left(-\frac{1}{2}\left(\frac{t - 5 \times 10^{-5}}{1.5 \times 10^{-5}}\right)^2\right). \quad (10)$$

At the inlet boundary, the compressible velocity perturbation, u_i', is set to zero, and a zero normal gradient condition is applied for the P, p', and ρ'. The solid boundaries are defined as slip walls for Eq. (1). For Eq. (2), the acoustically hard boundary, as defined in Eq. (9) is applied. At the far-field boundary, P is set to the ambient pressure p_∞, while p' and ρ' are set to zero. For both the INS velocity, U_i, and the perturbed velocity, u_i', a zero normal gradient condition is applied at the far-field boundary. The duct flow is initialized with ambient values. Specifically, the velocity and all perturbations are set to zero, P is set to p_∞, and the density is set to $\rho_0 = 1293 \text{ g/cm}^3$ at temperature $273K$. To provide a reference, the duct flow is also simulated by solving a fully compressible
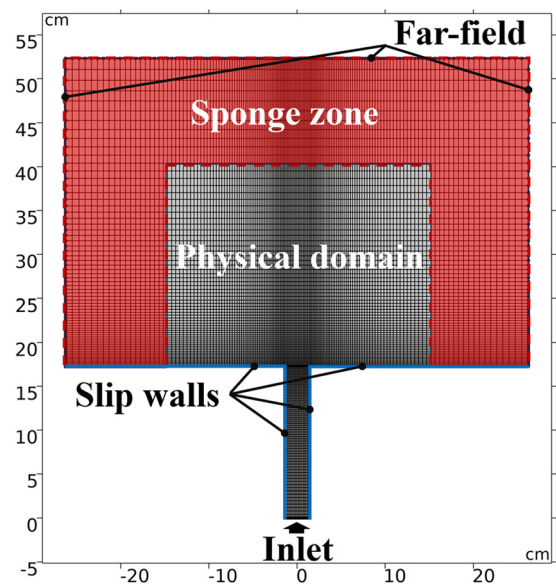

FIG. 4. Computational domain and mesh of the straight duct for the validation study.

Navier–Stokes equation (NS). At the inlet, the temperature and pressure are set to $273K$ and p_∞, respectively, allowing the density to vary in accordance with the velocity pulse. The slip boundary is applied on the walls. The ambient pressure, p_∞, is prescribed on the far-field boundary. Both the INS/LPCE and the NS equations are solved in the COMSOL Multiphysics, using quadratic elements and an implicit method. The time step is 1×10^{-6} s for the INS/LPCE solver, which is sufficiently small to track the sound wave in the duct. To ensure a stable solution, the time step is set to 5×10^{-8} s for the NS solver.

The velocity pulse, as defined in Eq. (10), triggers a pressure perturbation at the inlet boundary. The compressible pressure perturbation in the duct is shown in Fig. 5 by contours. The centerline of the duct is marked by the dashed line in Fig. 5(a). In Fig. 5(a), a compressed pressure wave propagates along the duct and scatters at the duct's end. During scatting, as shown in Fig. 5(b), a part of the pressure wave propagates upwards into the far field, while another part becomes an expanded pressure wave and propagates downwards into the duct. In Fig. 5(c), the expanded pressure wave bounces back from the end of the duct. The wave in the rectangular domain can pass through the boundary smoothly, demonstrating effectiveness of the sponge zone. The pressure signal along the centerline of the duct is shown in Fig. 6, and the results of the INS/LPCE solver and the NS solver are compared. It can be seen that the predictions of both solvers for the compressed and the expanded waves are in good agreement. By tracing the peak of the signal, the speed of the pressure wave is estimated to be 332 m/s, which is approximately the sound speed at $\sqrt{\gamma p_\infty / \rho_0}$ = 331 m/s. Thus, the prediction of the sound signal is reasonable, which validates the current INS/LPCE solver. This solver will be utilized to analyze the acoustic properties of various vocal fold vibrations.

IV. RESULTS AND DISCUSSION

The flow field is the source of the acoustic signal. In this section, the formation of sound source term is analyzed, followed by an
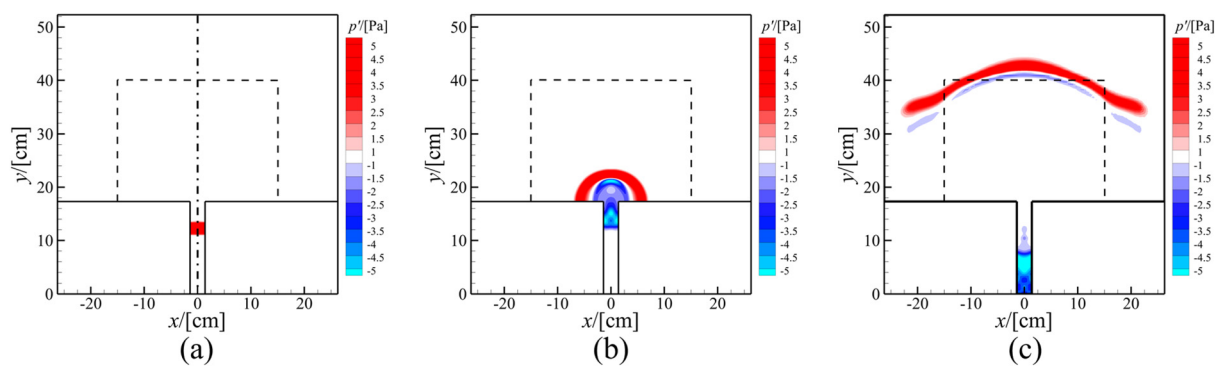
FIG. 5. Snapshots of the perturbed compressible pressure in the straight duct at time (a) $t = 4.5 \times 10^{-4}$ s, (b) $t = 7.5 \times 10^{-4}$ s, and (c) $t = 13.5 \times 10^{-4}$ s. Dashed lines mark the sponge zone, dash-dotted line in figure (a) marks the centerline.
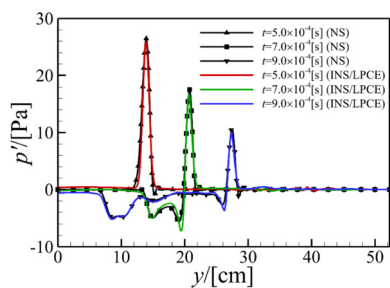

FIG. 6. Pressure signals along the centerline of the duct in the validation study.

investigation of the acoustic signal in trachea. An intensity ratio between flow and acoustic signals is examined to quantify the efficiency of voice production. The effects of the medial thickness and the frequency difference are discussed in these analyses.

As listed in Table I, the medial thickness and asymmetry of the vibration frequency are varied, and a total of 12 cases are considered, while the maximum flow rate q_{max} is the result from current modeling. The channel width D_0 is 2.2 cm. The minimum and maximum gaps,

TABLE I. Parameters of the vocal fold vibration in consideration.

No.	$FD = (f_B - f_A)/f_A$	M/cm	H	$q_{max}/[ml/(s \cdot cm)]$
Case 1	0.0%	0.13	1.9	1427
Case 2	0.0%	0.19	2.8	1400
Case 3	0.0%	0.28	4.2	1354
Case 4	5.0%	0.13	1.9	1419
Case 5	5.0%	0.19	2.8	1395
Case 6	5.0%	0.28	4.2	1353
Case 7	10%	0.13	1.9	1414
Case 8	10%	0.19	2.8	1371
Case 9	10%	0.28	4.2	1332
Case 10	−100%	0.13	1.9	1724
Case 11	−100%	0.19	2.8	1688
Case 12	−100%	0.28	4.2	1652

as defined in Eq. (4), are $D_{min} = 8 \times 10^{-2}$ cm and $D_{max} = 4.5 \times 10^{-1}$ cm, respectively. The frequency of fold-A is $f_A = 150$ Hz, which lies in physiological range of human phonation (Holmberg and Perkell, 1989). is used as the base frequency. The corresponding base period is $T_0 = 1/f_A$. The medial thickness, M, varies from 1.3×10^{-1} cm to 2.8×10^{-1} cm, and the corresponding coefficient of thickness, H, as defined in Eq. (5), is included in Table I. The range of M is selected based on simulations from Chen *et al.* (2020). The frequency of fold-B is 5% to 10% higher than that of the fold-A which is consistent with the measurements obtained from a high-speed video endoscopy by Li *et al.* (2021). The case of UVFP is approximated by fixing fold-B and is included in the last three rows of Table I. All cases are simulated on a structured mesh with 14 400 (45×320) grid cells. The grid spacing is 1.2×10^{-2} cm near the vocal fold and is stretched to 0.27 cm near the outlet. The time step size is 5×10^{-5} s. For statistical sampling, cases 1–3 and 7–12 are simulated for $20T_0$, while cases 4–6 are simulated for $50T_0$. These simulation durations are used to ensure that the results have reached a quasi-periodic state. Details about the grid sensitivity and the modeled asymmetry vocal fold motion are provided in the subsections 2 and 3 of Appendix, respectively.

A. Flow field and sound source term

The detailed flow field data are available from the incompressible flow simulation (see the supplementary material S1). Here we focus on the characteristics of the flow rate and those of DP/Dt, which is the only source term of compressible perturbations in Eq. (2).

The normalized flow rate $x/D_0 = 1$ is plotted in Fig. 7 to show its waveform. q_{max} is the maximum flow rate whose value is listed in Table I for each case. The available experimental data from Holmberg and Perkell (1989), in which the flow rate was measured at the mouth of a male speaker for a high pitch is shown here to compare with a symmetric (healthy) case. Since similar results are observed within case groups 1–3, 4–6, 7–9, and 10–12, respectively, only cases 2, 5, 8, 11 are shown for comparison. In Fig. 7(a), t/T_0 in case 2 (a symmetric setup) shows a regular pattern and is in good agreement with the experimental result. In Figs. 7(b) and 7(c), q/q_{max} appears to be modulated in oscillation magnitude, as marked by dashed lines. In Fig. 7(b), the time between two minimum oscillations is $20T_0$, corresponding to modulation frequency $5\%f_A$, which is exactly the frequency difference between the two folds. A similar result is observed in Fig. 7(c), in which
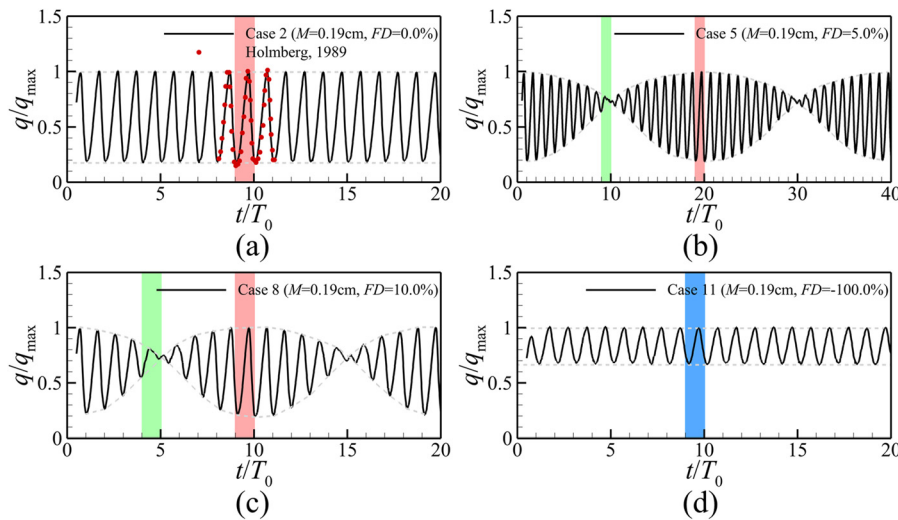
FIG. 7. Normalized flow rate for (a) case 2, (b) case 5, (c) case 8, (d) case 11. The oscillation magnitude is marked by dashed lines. A few characteristic cycles of the flow rate marked in color will be used for the analysis of acoustic source term; red: maximum or incited oscillation; green: minimal or suppressed oscillations; blue: UVFP.

the time between two minimum oscillations is $10T_0$, corresponding to modulation frequency $10\%f_A$. These results indicate that the flow rate is modulated at the frequency difference. In Fig. 7(d) for the UVFP, the pattern of q/q_{max} is regular (since the modulation frequency is $100\%f_A$), but the magnitude of oscillation is only one-third of that in Fig. 7(a). Because of the modulation, the average magnitude of oscillation in Figs. 7(b) and 7(c) is also much smaller compared to that in Fig. 7(a), and it is reduced to 65% in Fig. 7(b) and 69% in Fig. 7(c).

The distributions of DP/Dt in typical cycles in Fig. 7 are plotted in Figs. 8–10, for comparison. According to oscillation magnitudes, three kinds of cycles are distinguished in Fig. 7. They are the cycle with maximum oscillation magnitude, denoted by $C_{incited}$, the cycle with

minimum oscillation magnitude, denoted by $C_{suppressed}$, and the cycle in the case of the UVFP, denoted by C_{UVFP}. In a cycle of the flow rate oscillation, the development of DP/Dt in vocal fold is similar among the cases where the oscillation magnitudes of q/q_{max} are similar. Thus, only the field DP/Dt in typical cycles of the flow rate oscillation are compared. The field at $85\%T_0$ of the cycle $C_{incited}$, $C_{suppressed}$ and C_{UVFP} are shown in Figs. 8–10, respectively, at which the intensity of DP/Dt is significant. For the cycle $C_{incited}$, the results in cases 2, 5, 8 are found similar to each other, thus only the result in case 2 is shown in Fig. 8 (Multimedia view). In Fig. 8(a), the variation of DP/Dt mainly occurs in the range of $x/D_0 \in [-M/D_0, M/D_0]$, which is marked by a gray band. In this range, the two sides of the medial surface form a
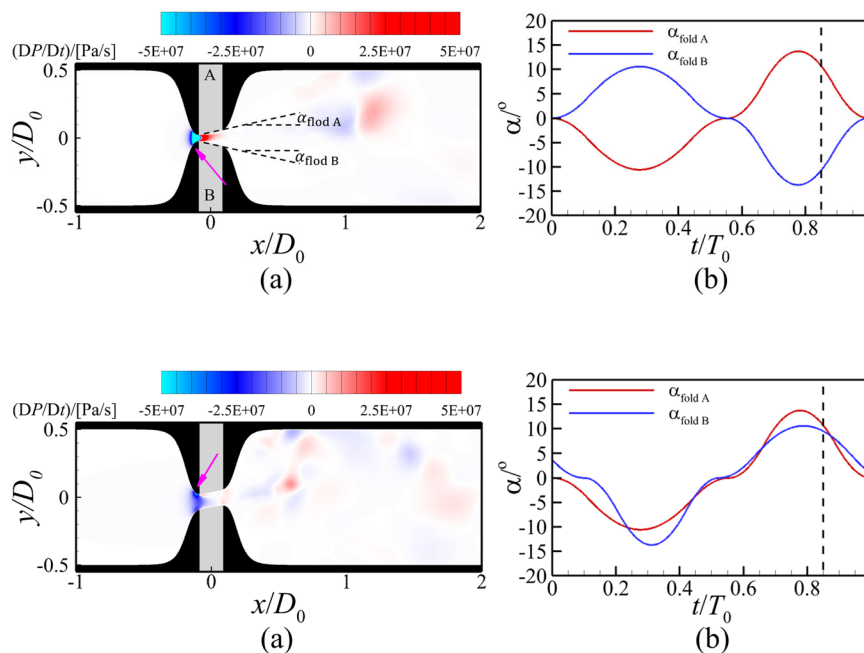
FIG. 8. Distribution of DP/Dt at $85\%T_0$ of the typical cycle $C_{incited}$ in case 2 ($M = 0.19$ cm, $FD = 0.0\%$). (a) Contour of DP/Dt, (b) Tilt angle of the medial surface. Similar distribution is observed in cases 2, 5, 8. In figure (a), the region of the medial surface is marked in gray. The concentration of DP/Dt is marked by a pink arrow. Dashed line in figure (b) marks $85\% T_0$. Multimedia available online.
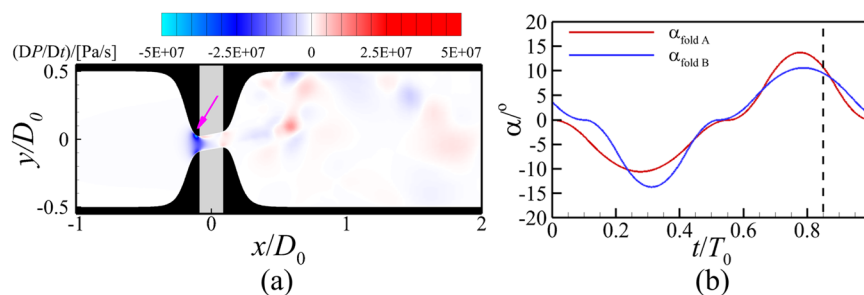

FIG. 9. Distribution of DP/Dt at $85\% T_0$ of the typical cycle $C_{suppressed}$ in case 5 ($M = 0.19$ cm, $FD = 5.0\%$). (a) Contour of DP/Dt, (b) Tilt angle of the medial surface. Similar distributions is observed in cases 5 and 8. The meaning of gray bar, the pink arrow and the dashed line are the same as those in Fig. 8. Multimedia available online.
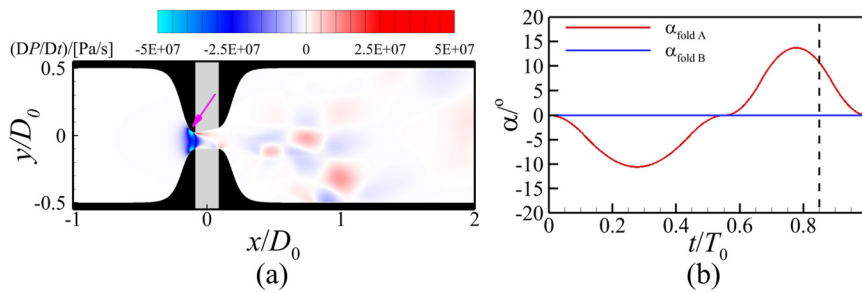
FIG. 10. Distribution of DP/Dt at 85% T_0 of the typical cycle C_{UVFP} in case 11 ($M = 0.19$ cm,　$FD = -100.0\%$).　(a) Contour of DP/Dt, (b) Tilt angle of the medial surface. The meaning of gray bar, the pink arrow and the dashed line are the same as those in Fig. 8. Multimedia available online.

divergent channel, in which DP/Dt exhibits an opposite-polarity source pair near the entrance, which is similar to a dipole source (Stewart and William, 2017). In Fig. 8(b), α of fold A and B are in opposite sign. For the cycle $C_{suppressed}$, the results in cases 5 and 8 are found similar to each other, thus only the result in case 5 is shown in Fig. 9 (Multimedia view). In Fig. 9(a), the medial surface exhibits a shape of an inclined parallel channel. In Fig. 9(b), α of fold A and B are approaching each other. In Fig. 10(a) (Multimedia view), the lower-side medial surface is horizontal, but the other side is divergent, resulting in a positive α as marked by dashed line in Fig. 10(b). Both in Figs. 9(a) and 10(a), DP/Dt concentrates on the immediate entrance of the glottal channel, as marked by pink arrows in the figures, and the feature of the opposite-polarity source pair becomes indistinct. The DP/Dt is a source term in form of material derivative. In the studies by Schoder et al. (2023) and Tieber et al. (2024), the magnitude of this type of source term is relatively low, due to cancelation between its temporal and convection components under the assumption of Taylor's hypothesis. However, in the case of the vocal fold, the pressure gradient can be significant, thereby diminishing the validity of Taylor's hypothesis (Geng et al., 2015). As DP/Dt is the only source term of Eq. (2), the opposite-polarity source pair of DP/Dt is important for the production of compressible perturbations. This result also suggests that the pressure variations experienced by the fluid elements in the channel serve as an important source of these perturbations. From the current study, the symmetric divergent channel is instrumental for the formation of the opposite-polarity source pair.

The temporal-spatial features of DP/Dt and p' along the centerline of the domain are shown in Fig. 11 by an x-t diagram. Despite the differences in the case setup, the results for each kind of the cycle (i.e., $C_{incited}$, $C_{suppressed}$ and C_{UVFP}) are found similar among the cases (see the supplementary material S2). Thus, only results of $C_{incited}$ in case 2, $C_{suppressed}$ in case 5, and C_{UVFP} in case 11 are shown as representative cycles for comparison. In Fig. 11(a) for the cycle $C_{incited}$, the opposite-polarity source pair of DP/Dt is formed at the entrance of the glottal channel and during a short time window when the vocal fold is closing. The corresponding pressure signal p' in the cycle $C_{incited}$ is concentrated in between the opposite-polarity source pair of DP/Dt, and its peak time coincides with the formation of the opposite-polarity source pair. This confirms that the opposite polarity paired sources of DP/Dt is responsible for the production of p'. In comparison with the cycle $C_{incited}$, the opposite-polarity source pair for the cycle $C_{suppressed}$ in Fig. 11(b) is not clear, and the spatial distribution of DP/Dt is nearly steady over time in the shown cycle. The corresponding signal of p' has low intensity. For the cycle C_{UVFP} in Fig. 11(c), the opposite-polarity source pair again is not clear, but the temporal change is significant. Correspondingly, p' has a greater intensity as compared to

that of the cycle $C_{suppressed}$. These results suggest that the frequency difference diminishes the opposite-polarity source pair of DP/Dt, thereby reducing the intensity of p'.

Along centerline of the vocal fold, DP/Dt of cases 1–3 at 85%T_0 of the cycle $C_{incited}$ are compared in Fig. 12, to show the effect of medial thickness. It is clear to see that, from cases 1 to 3, the extrema shift upstream with the distance between them decreasing and their absolute value increasing. In such a way the opposite-polarity source pair gradually becomes more compact and stronger. Although not presented here, the same trend is observed in case groups: 4–6, 7–9, and 10–12. Therefore, the increase in medial thickness affects both the distribution and the strength of the source term, resulting in a strong opposite-polarity source pair of DP/Dt.

In summary, the opposite-polarity source pair of DP/Dt occurs at the entrance of the divergent channel which is formed along the medial surface. Such a paired distribution of sources is responsible for the sound generation. The frequency difference lowers the intensity of the paired sources by reducing the divergent angle of the channel. The increase in medial thickness makes the distribution of the paired sources more compact and stronger.

B. Acoustic signal in trachea

The perturbed compressible pressure p' contains the waves that can radiate into the far field, as discussed in the subsection 4 of Appendix, and thus is employed to represent the acoustic signal in the trachea. The cases of different frequency differences are compared in Fig. 13, where p' is extracted at $(x/D_0, y/D_0) = (7, 0)$. Since similar results are observed within case groups 1–3, 4–6, 7–9, and 10–12, respectively, only cases 2, 5, 8, 11 are shown for comparison (see the supplementary material S3). It is known that the superposition of two tuning signals generates an amplitude-modulated signal when the frequencies of these signals are close to each other (Crawford, 1968). Specifically, for two tuning signals of frequency f_1 and f_2, a beat frequency at $f_{beat} = |f_1 - f_2|$ will emerge and is the repetition rate of the loud instant in the superposed signals (Crawford, 1968).

In Fig. 13(a), the amplitude is regular for case 2, indicating stable loudness of the signal. In Fig. 13(b) for case 5, the signal exhibits amplitude modulation. The time between two loud instants is $20T_0$, corresponding to the beat frequency of $5\%f_A$ and is consistent with the frequency difference in case 5. Similarly, the beat frequency in Fig. 13(c) is around $1/(10T_0) = 10\%f_A$, matching the frequency difference in case 8. Therefore, the acoustic signal is found to be modulated at the frequency difference. It is also clear that the amplitudes in Figs. 13(b) and 13(c) become low between loud instants, which necessarily reduces the audibility of the signal. In Fig. 13(d) for UVFP,
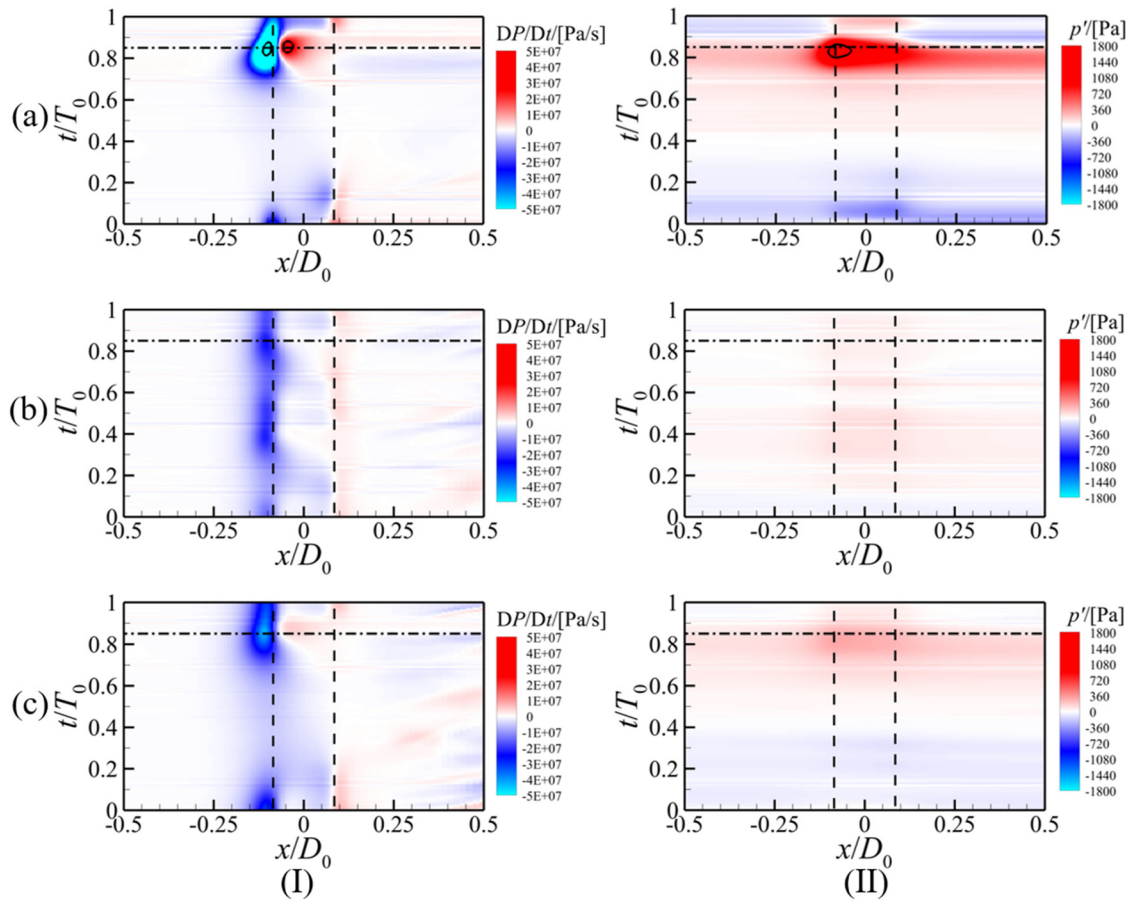
FIG. 11. The *x-t* diagram of (I) DP/Dt and (II) p' along the centerline of the vocal fold. Rows (a), (b), and (c) are results of the cycle C_{incited} in case 2 ($M=0.19\,\text{cm}$, $FD=0.0\%$), $C_{\text{suppressed}}$ in case 5 ($M=0.19\,\text{cm}$, $FD=5.0\%$), and C_{UVFP} in case 11 ($M=0.19\,\text{cm}$, $FD=-100.0\%$), respectively. Contour lines in row (a) mark 80% of the extrema. Vertical dashed lines mark the medial surface region of $\pm M/D_0$. The horizontal dash-dotted line marks 85% T_0 of the cycle.

although the amplitude is regular, the average magnitude of the signal is about 32% of that in Fig. 13(a). Thus, the UVFP significantly decreases the amplitude of sound. Because of the amplitude modulation, the average magnitude of the signal in Fig. 13(b) and 13(c) is also much smaller compared to that in Fig. 13(a), and it is reduced to 56% in Fig. 13(b) and 59% in Fig. 13(c).

Power spectral density (PSD) of the signal in cases with different frequency differences are compared in Fig. 14. The PSD is estimated by fast Fourier transformation. In Fig. 14, spectrum of case 2 has a peak at base frequency $f_A = 150$ Hz, and two peaks at harmonic frequencies 300 and 450 Hz, respectively. In cases with frequency difference, the PSD peaks are still around the base and harmonic frequencies. However, they spread and form a wide sideband. For case 11, the PSD peak value is lower than other cases, showing the UVFP suppressing all the frequency components.

The acoustic signals in cases of different medial thicknesses are compared in Fig. 15, where cases 1–3 are shown. From the waveform, a relatively quiet period is observed from 20% T_0 to 65% T_0 for these cases, which is followed by a large fluctuation period from 65% T_0 to T_0. All three cases show a similar pattern and reach a peak at 80% T_0, when the vocal fold is closing. The magnitude of the signal increases when the

medial thickness is raised from cases 1 to 3. Therefore, an increase in medial thickness tends to raise the peak value of the acoustic signal. Figure 16 shows the PSD of these signals. The PSD peaks at harmonic frequencies are higher than those at base frequency. At 450 Hz, the PSD peak in the case with $M = 0.28$ is nearly twice as high as that in the case with $M = 0.13$, indicating the increase in medial thickness enhances the harmonic component in the signal. These results agree with the tissue experiments reported by Zhang and Chhetri (2019). Their study also showed that a thick medial surface improves voice production and promotes higher-order harmonic excitation.

C. Intensity ratio between p' and P_r

The intensity ratio between the acoustic signal and the incompressible relative pressure P_r is defined by the following equations

$$R = \frac{\displaystyle\int_{\Omega} \left(p'(x,0,t)\right)^2 dxdt}{\displaystyle\int_{\Omega} \left(P_r(x,0,t)\right)^2 dxdt}, \tag{11}$$
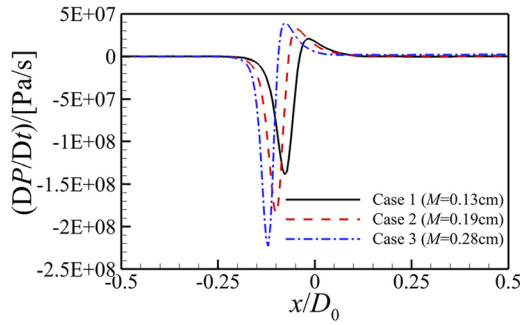
FIG. 12. Comparison of the instantaneous source term, DP/Dt, along the centerline of the vocal fold, at 85% T_0 of the cycle $C_{incited}$ for cases with different medial thickness.

$$I(p') = \frac{1}{S}\int_{\Omega}\left(p'(x,0,t)\right)^2 dxdt, \quad (12)$$

$$I(P_r) = \frac{1}{S}\int_{\Omega}\left(P_r(x,0,t)\right)^2 dxdt, \quad (13)$$

where $\Omega = \left\{(x,t)|(x/D_0, t/T_0) \in [1,9] \times [5,20]\right\}$ and $I(P_r)$ are the intensity of p' and P_r, respectively. $\Omega = \left\{(x,t)|(x/D_0, t/T_0) \in [1,9] \times [5,20]\right\}$ is the integral domain representing the centerline downstream the vocal fold. S is the area of Ω. R is the intensity ratio, which quantifies the glottal efficiency (Titze, 1992). The results are listed in Table II and plotted in Fig. 17. It is found that R is generally less than 25.5%, showing that the acoustic signal is not a dominant part of the pressure oscillation in the flow. Figure 17(a) shows that the intensity ratios in cases of asymmetric vibration are lower than those in cases of symmetric vibration. The ratio tends to increase as the medial thickness increases. In Fig. 17(a), R is nearly linear correlated with M/D_0. To quantify the sensitivity of R to M/D_0, the slope between them is estimated by linear regression. In cases with frequence difference varying from 0.0% to 10.0%, the average slope is 0.8. These results suggest that,
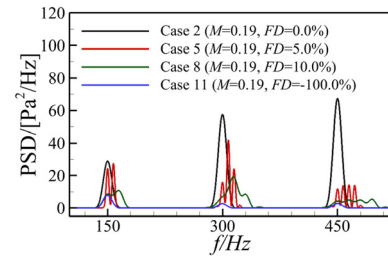

FIG. 14. Power spectral density of the signal p' for cases with different frequency differences. (a) Case 2, (b) case 5, (c) case 8, and (d) case 11.
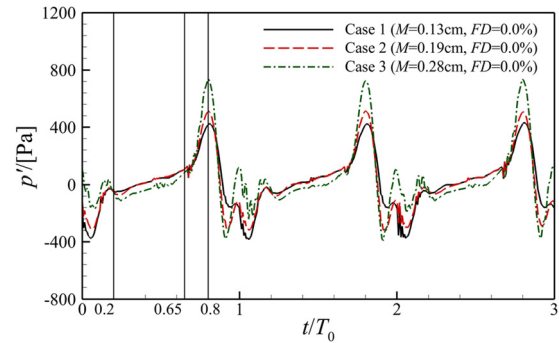

FIG. 15. Comparison of the signal for cases with different medial thicknesses. The signals are extracted at $(x/D_0, y/D_0) = (7,0)$, for three periods.

under the same driving pressure, a thick medial surface is able to improve the glottal efficiency. In cases with frequency difference equaling -100.0%, the intensity ratio is significantly lower than in the other cases, suggesting a largely reduced glottal efficiency in the UVFP. And for these cases, the slop reduces to 0.1, showing the effect of medial surface is suppressed. The effect of frequency difference on the intensity ratio is shown in Fig. 17(b). It can be seen that R in cases with 5%
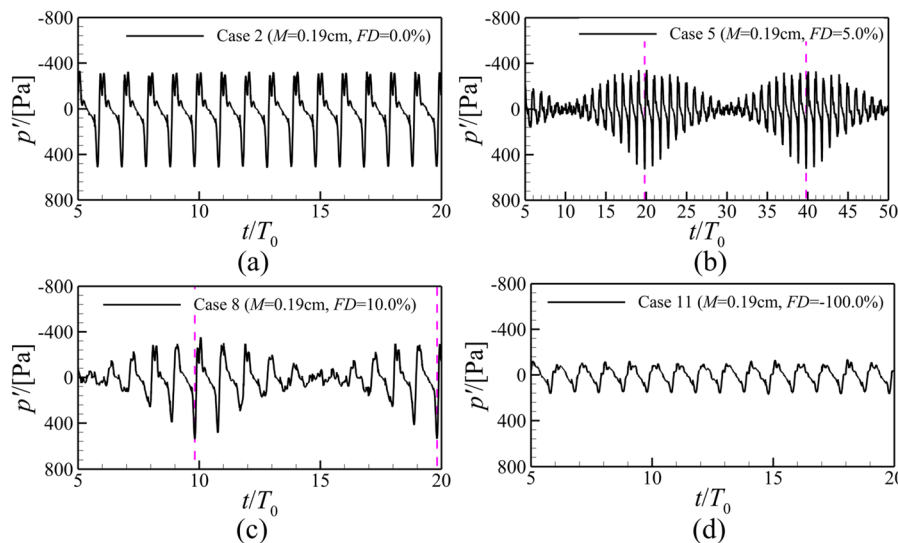

FIG. 13. Comparison of the acoustic signal p' for cases with different frequency differences. (a) Case 2, (b) case 5, (c) case 8, and (d) case 11. The signals are extracted at $(x/D_0, y/D_0) = (7,0)$. The vertical dashed lines mark instants of loud sound.
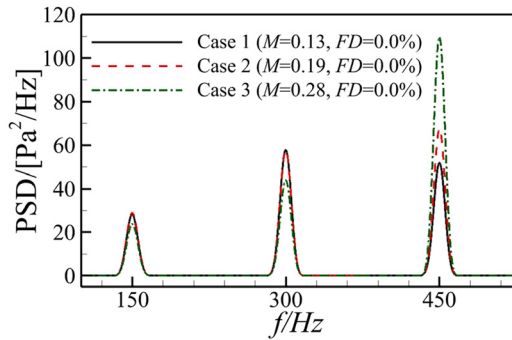
FIG. 16. Power spectral density of the signal p' for cases with different medial thicknesses.

TABLE II. Intensity ratios between P_r and p'.

ID	Intensity of p' (Pa^2)	Intensity of P_r (Pa^2)	R (%)
Case 1	3.15×10^4	1.71×10^5	18.4
Case 2	3.66×10^4	1.68×10^5	21.8
Case 3	4.60×10^4	1.81×10^5	25.5
Case 4	1.38×10^4	1.04×10^5	13.3
Case 5	1.48×10^4	1.01×10^5	14.7
Case 6	1.68×10^4	9.83×10^4	17.1
Case 7	1.59×10^4	1.16×10^5	13.7
Case 8	1.76×10^4	1.14×10^5	15.4
Case 9	2.13×10^4	1.10×10^5	19.4
Case 10	5.00×10^3	1.07×10^5	4.67
Case 11	5.14×10^3	1.01×10^5	5.10
Case 12	5.50×10^3	9.74×10^4	5.65

frequency difference are lower than those with 10% frequency difference. The variation of R is not proportional to the increase in frequency difference, which indicates a nonlinear effect of asymmetric vibration.

V. CONCLUSION

A flow-acoustic model is developed and validated to investigate the effect of asymmetric vocal fold vibration on sound production. The perturbed compressible pressure and the source term of LPCE are analyzed to shed light on the sound generation process. The main conclusions are summarized as follows:

(1) During asymmetry vocal fold motion, three typical cycles of the normalized flow rate are identified as the cycle with maximum oscillation magnitude, denoted by C_{incited}, the cycle with minimum oscillation magnitude, denoted by $C_{\text{suppressed}}$, and the cycle in case of the UVFP, denoted by C_{UVFP}. Each cycle shows a different distribution of the source term DP/Dt, and in the cycle C_{incited}, DP/Dt exhibits an opposite-polarity source pair that is primarily responsible for the production of the perturbed compressible pressure, p'.

(2) The opposite-polarity source pair of DP/Dt occurs at the entrance of the glottal channel and is more pronounced for a divergent channel formed by the medial surfaces when the vocal fold is closing. The frequency difference lowers the intensity of the opposite-polarity source pair by reducing the divergent angle of the channel. The increase in the medial thickness makes the distribution of the opposite-polarity source pair more compact and also stronger in magnitude.

(3) The frequency difference reduces the acoustic signal through amplitude modulation, therefore leading to voice problems. The increase in medial thickness tends to enhance the intensity ratio between the acoustic signal and the hydrodynamic pressure variations and thus increases the glottal efficiency, while the frequency difference, especially the UVFP, reduces the intensity ratio and the glottal efficiency.

VI. CURRENT SHORTCOMINGS

The current model has the following shortcomings, which will be addressed in future work. (1) The model is in a two-dimensional space, assuming the vocal fold motion is synchronized in out-of-plane direction. However, the three-dimensional motion of the vocal fold can be desynchronized in the out-of-plane direction, which may introduce additional modulation to the acoustic signal. (2) The vocal fold motion is prescribed, and as a result, flow-induced tissue deformation is not considered. This simplification may affect the magnitude of the acoustic signal. (3) The acoustic properties are analyzed based on p'. However, as discussed in the subsection 4 of Appendix, p' generally consists of multiple components related to compressibility. Among them, the radiative component is more suitable for acoustic analysis.

SUPPLEMENTARY MATERIAL

See the supplementary material, which presents figures of the distribution of P_r, the temporal-spatial development of DP/Dt and the signal of p'.
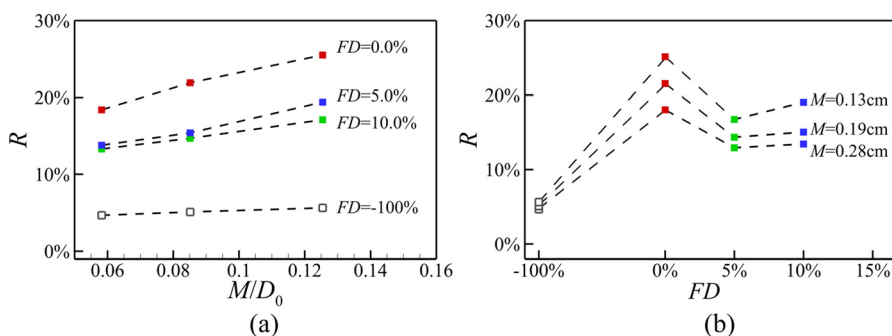

(a)

(b)

FIG. 17. Intensity ratio R. (a) The effect of the medial thickness M. Results with the same frequency difference (FD) are linked by dashed line. (b) The effect of the frequency difference $FD = (f_B - f_A)/f_A$. Results with the same medial thickness are linked by dashed line.

ACKNOWLEDGMENTS

AUTHOR DECLARATIONS
Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Qilin Liu: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Guofeng He:** Data curation (equal); Investigation (equal); Software (equal). **Lea Sayce:** Conceptualization (equal); Formal analysis (equal); Supervision (equal); Writing – original draft (equal); Writing – review & editing (equal). **Haoxiang Luo:** Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Supervision (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Zheng Li:** Conceptualization (equal); Data curation (equal); Formal analysis (equal); Funding acquisition (equal); Investigation (equal); Supervision (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

APPENDIX: SUPPLEMENTARY REMARKS ON MODELING
1. Numerical treatments for the LPCE

In the sponge zone a damping term, D_ϕ, and a convection term, C_ϕ, are added to the right and left sides of Eq. (2), respectively, to annihilate the waves. These terms can be written as follows (Freund, 1997)

$$D_\phi = -A_d \cdot \sigma \cdot (\phi - \phi_0), \tag{A1}$$

$$C_\phi = A_c \cdot \sigma \cdot \frac{\partial \phi}{\partial x_i}, \tag{A2}$$

where ϕ represent the primitive variables: ρ', u'_i, and p'. ϕ_0 is a target value of ϕ. A_d and A_c are the strength of damping and convection, respectively. σ is an exponential damping function

FIG. 19. Asymmetric motion of a pair of vocal folds. (a) The gap between two folds measured by high-speed video-endoscopy from an *in vivo* test (Li *et al.*, 2021), (b) the modeled gap at $x/D_0 = 0$ with frequency $f_B = (1 + 10\%)f_A$. Dashed and dash-dotted lines mark the in-phase and out-of-phase states, respectively.

$$\sigma = \frac{1}{1 + \exp\left(\frac{1}{\xi - 1} + \frac{1}{\xi}\right)}, \quad \xi = \frac{x - x_s}{x_e - x_s}, \tag{A3}$$

where x_s and x_e is the start and end of the zone. To absorb the waves, the ϕ_0 is set to zero for these primitive variables, and the damping strength $A_d = \rho_\infty c_\infty^2$ is applied to match the impedance at the boundary. The convection strength is set to $A_c = 1.05 c_\infty$. In this way, any perturbations in the sponge zone will be connected out of the computational domain. The D_ϕ is added for both the inlet and outlet sponge zones. The C_ϕ is added only for the outlet sponge zone. In the present model, the start position, x_s, of the convection zone is set 50% ahead of the damping zone to achieve a good non-reflection feature. To be compatible with the sponge zone, the variables ρ' and p' at the inflow boundary are set to zero. Their normal gradients are set to zero at the outflow boundary. The normal gradient of u'_i is maintained at zero at both the inflow and outflow boundaries to enable the convection term to take effect.

In addition to the boundary treatment, an artificial viscosity term is added on the left side of Eq. (2) to stabilize the solution. The term is written as

$$-c \frac{\partial^2 \phi}{\partial x_j \partial x_j}, \tag{A4}$$

where coefficient $c = \delta \cdot \Delta \cdot \beta$, the Δ is the scale of the grid cell, the β is the maximum speed within that grid cell. When the quadratic element is applied for the discretization, the $\delta = 0.25$ is suggested to stabilize the solution (COMSOL, 2023).
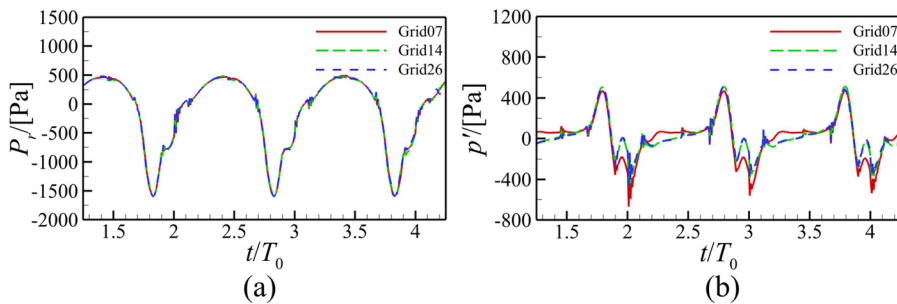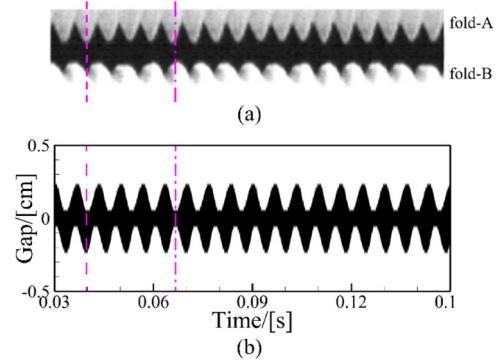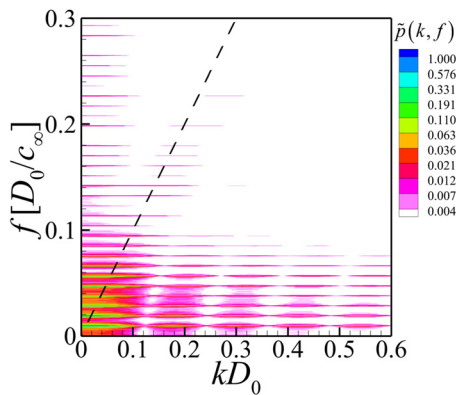

FIG. 18. Grid independence test for (a) incompressible pressure signal and (b) perturbed compressible pressure signal.

FIG. 20. The *k-f* diagram of the perturbed compressible pressure signal. The dashed line marks the slope c_∞. The color axis is in logarithmic scale.

2. Grid sensitivity

Simulations on grids with 0.7×10^4, 1.4×10^4, 2.6×10^4 cells are carried out for case 2 to analyze grid sensitivity. The three grids are denoted as Grid07, Grid14, and Grid26, respectively. At position $(x/D_0, y/D_0) = (0.5, 0)$, signals of relative pressure P_r and p' are shown in Fig. 18 for three periods of vocal fold motion. It can be seen from the plots that Grid26 does not significantly improve the results, and Grid14 is sufficient for the present study.

3. Asymmetry motion of vocal fold

The temporal variation of the gap between a pair of folds is visualized by kymograph as shown in Fig. 19. Figure 19(a) shows data from a previous *in-vivo* test using evoked rabbit phonation, measured by high-speed video endoscopy (Li *et al.*, 2021). From the test, both the in-phase and out-of-phase of the vocal fold motion can be observed, marked by dashed and dash-dotted lines, respectively. Figure 19(b) shows the modeled vocal fold motion of case 8, in which the frequency $f_B = (1 + 10\%)f_A$. The modeled motion displays a similar pattern to that in Fig. 19(a). This result indicates that the present model is able to represent the asymmetric motion of the vocal fold.

4. Components in perturbed compressible pressure

To visualize the components of the perturbed compressible pressure, the signal $p'(x, t)$ is transformed into the frequency-wavenumber domain by Fourier transformation, as follows:

$$\tilde{p}(k, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p'(x, t) \exp(-2\pi i(k \cdot x - f \cdot t)) dx dt, \quad \text{(A5)}$$

where i is the imaginary unit, k and f are the wavenumber and frequency of the signal, respectively. For a traveling wave, the k and f are related by the phase, $\varphi = 2\pi(k \cdot x - f \cdot t)$. The propagation speed of the wave can be estimated by letting $d\varphi = (k \cdot dx - f \cdot dt) = 0$, so the phase speed of the wave is calculated as

$$\nu_\phi = \frac{dx}{dt} = \frac{f}{k}. \quad \text{(A6)}$$

The ν_ϕ of the radiative wave should not be lower than the sound speed c_∞ at $\sqrt{\gamma p_\infty / \rho_\infty} = 354$ m/s. Thus, it is necessary for the radiative wave to satisfy the condition $|f| \geq c_\infty |k|$ (Lighthill, 1962). Otherwise, the perturbation wave is quickly dampened by the medium, so-called hydrodynamic wave (Liu and Lai, 2021; Mancinelli *et al.*, 2018).

Along the centerline, p' signal is transformed from the spatial-temporal domain $(x/D_0, t/T_0) \in [1, 9] \times [5, 20]$ into the frequency-wavenumber domain and the result of case 2 is shown in Fig. 20 by the *k-f* diagram. In Fig. 20, the sonic line is marked by the dashed line, along which the ratio $\nu_\phi = f/k$ is c_∞. In Fig. 20, the signal is a series of horizontal stripes in the *k-f* diagram. The dominant stripe is found at nondimensional frequency $f_A c_\infty / D_0 \approx 0.01$, which corresponds to the base frequency. The other stripes are the harmonic frequencies of the base frequency. Another observation is that the map consists of bands that are generally parallel to the sonic line, indicating the harmonic wavenumber generated from the Fourier analysis. In such a way, the signal distributes both above and below the sonic line. Therefore, the signal consists of both the radiative component and the hydrodynamic component.

5. Video files

The video linked to Fig. 8 shows the development of DP/Dt during the cycle C_{incited} for cases 2, 5, and 8.

The video linked to Fig. 9 shows the development of DP/Dt during the cycle $C_{\text{suppressed}}$ for cases 5 and 8.

The video linked to Fig. 10 shows the development of DP/Dt during the cycle C_{UVFP} of case 11.

REFERENCES

Bae, Y. and Moon, Y. J., "Computation of phonation aeroacoustics by an INS/PCE splitting method," Comput. Fluids **37**(10), 1332–1343 (2008).

Bae, Y. and Moon, Y. J., "Computation of bio-fluid sounds," Int. J. Aeroacoust. **10**(4), 475–496 (2011).

Bodaghi, D., Jiang, W., Xue, Q., and Zheng, X., "Effect of supraglottal acoustics on fluid–structure interaction during human voice production," J. Biomech. Eng. **143**(4), 041010 (2021).

Bodaghi, D., Xue, Q., Zheng, X., and Thomson, S., "Effect of subglottic stenosis on vocal fold vibration and voice production using fluid–structure–acoustics interaction simulation," Appl. Sci. **11**(3), 1221 (2021).

Chen, Y., Li, Z., Chang, S., Rousseau, B., and Luo, H., "A reduced-order flow model for vocal fold vibration: From idealized to subject-specific models," J. Fluids Struct. **94**, 102940 (2020).

Choi, H. and Moin, P., "Grid-point requirements for large eddy simulation: Chapman's estimates revisited," Phys. Fluids **24**(1), 011702 (2012).

Chung, H. R., Lee, Y., Reddy, N. K., Zhang, Z., and Chhetri, D. K., "Effects of thyroarytenoid activation induced vibratory asymmetry on voice acoustics and perception," Laryngoscope **134**(3), 1327–1332 (2024).

COMSOL, *COMSOL Multiphysics Reference Manual (Version 6.2)* (COMSOL, 2023).

Crawford, F. S., *Waves, Berkeley Physics Course* (McGraw-Hill, 1968), Vol. III.

Döllinger, M., Zhang, Z., Schoder, S., Šidlof, P., Tur, B., and Kniesburges, S., "Overview on state-of-the-art numerical modeling of the phonation process," Acta Acust. **7**, 25 (2023).

Enver, N., Azizli, E., Akbulut, S., Çadalli Tatar, E., Yelken, M. K., Öztürk, K., Coşkun, H., Birkent, A. H., Büyükatalay, Z. Ç., Özgürsoy, O. B., and Oğuz, H., "Inflammatory complications of vocal fold injection with hyaluronic acid: A multiinstitutional study," Turk. J. Med. Sci. **51**(2), 819–825 (2021).

Ewert, R. and Schröder, W., "Acoustic perturbation equations based on flow decomposition via source filtering," J. Comput. Phys. **188**(2), 365–398 (2003).

Falk, S., Kniesburges, S., Schoder, S., Jakubaß, B., Maurerlehner, P., Echternach, M., Kaltenbacher, M., and Döllinger, M., "3D-FV-FE aeroacoustic larynx model for investigation of functional based voice disorders," Front. Physiol. **12**, 616985 (2021).

Freund, J. B., "Proposed inflow/outflow boundary condition for direct computation of aerodynamic sound," AIAA J. **35**(4), 740–742 (1997).

Geng, C., He, G., Wang, Y., Xu, C., Lozano-Durán, A., and Wallace, J. M., "Taylor's hypothesis in turbulent channel flow considered using a transport equation analysis," Phys. Fluids **27**(2), 025111 (2015).

He, Y., Zhang, X., Chen, T., Li, Y., Deng, T., and He, Y., "An immersed boundary method for modeling fluid–solid–acoustic interactions involving dynamic structures," Phys. Fluids **36**(9), 096138 (2024).

He, Y., Zhang, X., Zhang, T., Geng, J., Liu, J., and Chen, X., "An improved hydrodynamic/acoustic splitting method for fluid–structure interaction feedback with elastic boundaries," Phys. Fluids **34**(2), 023606 (2022).

Henriques, D. P., Martins, R. H. G., and Cataneo, A. J. M., "Efficacy of injectable laryngoplasty with hyaluronic acid and/or calcium hydroxyapatite in the treatment of glottic incompetence. Systematic review and meta-analysis," J. Voice **39**(4), 945–953 (2023).

Holmberg, E. B. and Perkell, J. S., "Glottal airflow and transglottal air pressure measurements for male and female speakers in low, normal, and high pitch," J. Voice **3**(4), 294–305 (1989).

Hwang, S. and Moon, Y. J., "On the computation of low-subsonic turbulent pipe flow noise with a hybrid LES/LPCE method," Int. J. Aeronaut. Space Sci. **18**(1), 48–55 (2017).

Jo, Y., Ra, H., Moon, Y. J., and Döllinger, M., "Three-dimensional computation of flow and sound for human hemilarynx," Comput. Fluids **134–135**(1), 41–50 (2016).

Kaltenbacher, M., *Computational Acoustics* (Springer International Publishing, 2018), Vol. 579.

Kaltenbacher, M., Hüppe, A., Grabinger, J., and Wohlmuth, B., "Modeling and finite element formulation for acoustic problems including rotating domains," AIAA J. **54**(12), 3768–3777 (2016).

Konjević, B., Košec, A., Peček, M., and Stevanović, S., "Management of implant extrusion following thyroplasty for vocal fold paralysis: A CASE report," ORL **86**(1), 50–53 (2024).

Kosako, S., Hiramatsu, M., Fujimoto, Y., and Tsuji, Y., "Downstream flow field structure in voice prosthesis and its effect on sound generation around the esophageal wall," Phys. Fluids **35**(2), 025114 (2023).

Lasota, M., Šidlof, P., Maurerlehner, P., Kaltenbacher, M., and Schoder, S., "Anisotropic minimum dissipation subgrid-scale model in hybrid aeroacoustic simulations of human phonation," J. Acoust. Soc. Am. **153**(2), 1052–1063 (2023).

Li, Z., "Subject-specific modeling of implant placement for type I thyroplasty surgery," Ann. Biomed. Eng. **51**(10), 2182–2191 (2023).

Li, Z., Wilson, A., Sayce, L., Avhad, A., Rousseau, B., and Luo, H., "Numerical and experimental investigations on vocal fold approximation in healthy and simulated unilateral vocal fold paralysis," Appl. Sci. **11**(4), 1817–1815 (2021).

Lighthill, M. J., "The Bakerian Lecture, 1961 Sound generated aerodynamically," Proc. R. Soc. London, Ser. A **267**(1329), 147–182 (1962).

Liu, Q. and Lai, H., "Flow developing properties of a compressible parallel jet," AIP Adv. **11**(2), 025214 (2021).

Mancinelli, M., Pagliaroli, T., Camussi, R., and Castelain, T., "On the hydrodynamic and acoustic nature of pressure proper orthogonal decomposition modes in the near field of a compressible jet," J. Fluid Mech. **836**, 998–1008 (2018).

Maurerlehner, P., Schoder, S., Tieber, J., Freidhager, C., Steiner, H., Brenn, G., Schäfer, K.-H., Ennemoser, A., and Kaltenbacher, M., "Aeroacoustic formulations for confined flows based on incompressible flow data," Acta Acust. **6**, 45 (2022).

McCollum, I., Badr, D., Throop, A., and Zakerzadeh, R., "Biotransport in human phonation: Porous vocal fold tissue and fluid–structure interaction," Phys. Fluids **35**(12), 121909 (2023a).

McCollum, I., Throop, A., Badr, D., and Zakerzadeh, R., "Gender in human phonation: Fluid–structure interaction and vocal fold morphology," Phys. Fluids **35**(4), 041907 (2023b).

Michaud-Dorko, J., Farbos De Luzan, C., Dion, G. R., Gutmark, E., and Oren, L., "Comparison of aerodynamic and elastic properties in tissue and synthetic models of vocal fold vibrations," Bioengineering **11**(8), 834 (2024).

Nerrière, E., Vercambre, M.-N., Gilbert, F., and Kovess-Masféty, V., "Voice disorders and mental health in teachers: A cross-sectional nationwide study," BMC Public Health **9**(1), 370 (2009).

Pickup, B. A. and Thomson, S. L., "Influence of asymmetric stiffness on the structural and aerodynamic response of synthetic vocal fold models," J. Biomech. **42**(14), 2219–2219 (2009).

Schoder, S., "PCWE for FSAI – Derivation of scalar wave equations for fluid-structure-acoustics interaction of low Mach number flows," arXiv:2211.07490 (2025).

Schoder, S., Spieser, É., Vincent, H., Bogey, C., and Bailly, C., "Acoustic modeling using the aeroacoustic wave equation based on Pierce's operator," AIAA J. **61**(9), 4008–4017 (2023).

Schoder, S., Weitz, M., Maurerlehner, P., Hauser, A., Falk, S., Kniesburges, S., Döllinger, M., and Kaltenbacher, M., "Hybrid aeroacoustic approach for the efficient numerical simulation of human phonation," J. Acoust. Soc. Am. **147**(2), 1179–1194 (2020).

Seo, J. H. and Mittal, R., "A high-order immersed boundary method for acoustic wave scattering and low-Mach number flow-induced sound in complex geometries," J. Comput. Phys. **230**(4), 1000–1019 (2011).

Seo, J. H. and Moon, Y. J., "Linearized perturbed compressible equations for low Mach number aeroacoustics," J. Comput. Phys. **218**(2), 702–719 (2006).

Šidlof, P., Zörner, S., and Hüppe, A., "A hybrid approach to the computational aeroacoustics of human voice production," Biomech. Model. Mechanobiol. **14**(3), 473–488 (2015).

Stewart, G. and William, D., *Aeroacoustics of Low Mach Number Flows Fundamentals, Analysis, and Measurement* (Academic Press, 2017).

Tam, C. K. W., Viswanathan, K., Ahuja, K. K., and Panda, J., "The sources of jet noise: Experimental evidence," J. Fluid Mech. **615**, 253–292 (2008).

Tieber, J., Steiner, H., Maurerlehner, P., Schoder, S., Schäfer, K., Ennemoser, A., Kaltenbacher, M., and Brenn, G., "Hybrid aeroacoustic investigation of turbulent 90° pipe bend flow with source terms from large-eddy simulation," Comput. Fluids **279**, 106323 (2024).

Titze, I. R., "Vocal efficiency," J. Voice **6**(2), 135–138 (1992).

Woo, P., Pearl, A. W., Hsiung, M.-W., and Som, P., "Failed medialization laryngoplasty: Management by revision surgery," Otolaryngol.-Head. Neck Surg. **124**(6), 615–621 (2001).

Xu, H. H. A. and Yang, X. I. A., "Treatment of unphysical numerical oscillations via local grid refinement," Phys. Fluids **33**(7), 077104 (2021).

Xue, Q., Mittal, R., Zheng, X., and Bielamowicz, S., "A computational study of the effect of vocal-fold asymmetry on phonation," J. Acoust. Soc. Am. **128**(2), 818–827 (2010).

Zhang, Z., "Effect of vocal fold stiffness on voice production in a three-dimensional body-cover phonation model," J. Acoust. Soc. Am. **142**(4), 2311–2321 (2017).

Zhang, Z., "Vocal instabilities in a three-dimensional body-cover phonation model," J. Acoust. Soc. Am. **144**(3), 1216–1230 (2018).

Zhang, Z., "Contribution of undesired medial surface shape to suboptimal voice outcome after medialization laryngoplasty," J. Voice **38**(5), 1220–1226 (2024).

Zhang, Z. and Chhetri, D. K., "Effect of changes in medial surface shape on voice production in excised human larynges," J. Acoust. Soc. Am. **146**(5), EL412–417 (2019).

Zhang, Z., Wu, L., Gray, R., and Chhetri, D. K., "Three-dimensional vocal fold structural change due to implant insertion in medialization laryngoplasty," PLoS One **15**(1), e0228464–11 (2020).

Zhao, W., Zhang, C., Frankel, S. H., and Mongeau, L., "Computational aeroacoustics of phonation, Part I: Computational methods and sound generation mechanisms," J. Acoust. Soc. Am. **112**(5), 2134–2146 (2002).

25 August 2025 14:22:32