DYNAMIC CHARACTERIZATION OF VOCAL FOLD VIBRATIONS

A Dissertation
Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in
The Department of Mechanical Engineering

by
Zhenyi Wei
B.S., Zhejiang University, China, 2001
M.S., Zhejiang University, China, 2004
August 2012
Acknowledgments

I would like to express my deepest gratitude to my advisor, Dr. Marcio de Queiroz, for his patient guidance and continuous support during my graduate years. His expertise, insight, and quest for perfection in research greatly impressed me and added great value to my research experience. Without his guidance, I would have never finished this dissertation. I would like to thank my committee member, Dr. Melda Kunduk, who guided me in this specialized medical area. Her knowledge in voice disorders and experience in clinical practice greatly help me understand the nature of vocal folds. I would also like to thank my collaborator, Jing Chen, for her wonderful work in pre-processing the raw data captured from the high-speed digital imaging system.

Gratitude is extended to all my friends, especially to An Wu and Ziliang Cai, for their help over the years. I would like to thank my parents and brothers for their support and encouragement. Last but not least, I would like to thank my wife, Jiongjiong for standing beside me.
Table of Contents

Acknowledgments ........................................................................................................ ii

List of Tables ............................................................................................................... v

List of Figures .............................................................................................................. vi

Abstract ......................................................................................................................... x

Chapter 1: Introduction ................................................................................................. 1
  1.1 Motivation ............................................................................................................... 1
  1.2 Vocal Fold Anatomy ............................................................................................. 3
  1.3 Voice Production ................................................................................................. 6
  1.4 Voice Disorders ................................................................................................... 7
  1.5 Vocal Fold Examination ....................................................................................... 8
  1.6 Scope of Work ....................................................................................................... 10

Chapter 2: Biomechanical Models of Vocal Folds ......................................................... 12
  2.1 One-Mass Model .................................................................................................. 12
  2.2 Two-Mass Model ............................................................................................... 13
  2.3 Multi-Mass Models ............................................................................................. 18
  2.4 Asymmetric Models ......................................................................................... 19
  2.5 Estimation Approaches for Model Parameters .................................................. 21

Chapter 3: Proposed Biomechanical Model ................................................................. 22
  3.1 Basic Nomenclature ............................................................................................. 22
  3.2 Model Description ............................................................................................... 23

Chapter 4: Tuning of Model Parameters ..................................................................... 29
  4.1 Manual Coarse-Tuning Procedure ..................................................................... 30
  4.2 Automatic Fine-Tuning Procedure ...................................................................... 32

Chapter 5: Results ......................................................................................................... 35
  5.1 Experimental System ......................................................................................... 35
  5.2 Data Processing .................................................................................................. 36
    5.2.1 Corrections for FFT Calculation .................................................................. 37
    5.2.2 Alignment ................................................................................................. 38
  5.3 Analysis of Results ............................................................................................. 40
    5.3.1 Comparison of Coarse and Fine Tuning ...................................................... 41
    5.3.2 Discussion ............................................................................................... 42

Chapter 6: Conclusion and Recommendations ........................................................... 48
List of Tables

5.1 Asymmetry parameters. ............................................ 46
A.1 Parameters for healthy subjects. ............................. 54
A.2 Parameters for subjects with polyp. ....................... 55
A.3 Parameters for subjects with nodules. ..................... 56
A.4 Parameters for subjects with unilateral paralysis. ........ 57
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Larynx structure and position.</td>
<td>4</td>
</tr>
<tr>
<td>1.2</td>
<td>Superior view of vocal folds. Left: supporting structure; right: superior view.</td>
<td>5</td>
</tr>
<tr>
<td>1.3</td>
<td>Left: frontal view of the left larynx. Right: enlarged view of the coronal section of the left vocal fold.</td>
<td>6</td>
</tr>
<tr>
<td>1.4</td>
<td>Schematic of a high-speed imaging system.</td>
<td>9</td>
</tr>
<tr>
<td>1.5</td>
<td>A typical glottal cycle of vibration of normal (healthy) vocal folds captured by a HSDI system.</td>
<td>10</td>
</tr>
<tr>
<td>1.6</td>
<td>Block diagram of the dynamic characterization scheme for the vocal fold vibration.</td>
<td>11</td>
</tr>
<tr>
<td>2.1</td>
<td>Schematic of the one-mass model.</td>
<td>13</td>
</tr>
<tr>
<td>2.2</td>
<td>Schematic of the two-mass model.</td>
<td>15</td>
</tr>
<tr>
<td>2.3</td>
<td>A cycle of the vocal fold vibration.</td>
<td>16</td>
</tr>
<tr>
<td>2.4</td>
<td>Schematic of the three-mass model.</td>
<td>19</td>
</tr>
<tr>
<td>3.1</td>
<td>Side view of the proposed multi-mass model.</td>
<td>24</td>
</tr>
<tr>
<td>3.2</td>
<td>Top view of the proposed multi-mass model. Only upper masses of cover layer are shown.</td>
<td>25</td>
</tr>
<tr>
<td>3.3</td>
<td>3-D view of the proposed multi-mass model.</td>
<td>26</td>
</tr>
<tr>
<td>3.4</td>
<td>Simulation results of symmetric vocal folds using the proposed multi-mass model.</td>
<td>28</td>
</tr>
<tr>
<td>4.1</td>
<td>Maximum and minimum displacements in cycle ( j ).</td>
<td>33</td>
</tr>
<tr>
<td>4.2</td>
<td>Flowchart of fine tuning procedure.</td>
<td>34</td>
</tr>
<tr>
<td>5.1</td>
<td>Geometry of the glottal contour, glottal axis, and displacement of a contour point.</td>
<td>36</td>
</tr>
<tr>
<td>5.2</td>
<td>Contour of the glottis vibration during one cycle.</td>
<td>37</td>
</tr>
</tbody>
</table>
5.3 IFFT method for frequency correction. .............................. 39
5.4 Parabolic interpolation for time delay estimation. ............... 40
5.5 (a) Experimental and simulation data before alignment. (b) Experimental and simulation data after alignment. ................. 41
5.6 Plot of $Q_{l/r}$ versus $Q_{a/p}$ and classification of asymmetry plane. . . 47
B.1 Manual coarse tuning for subject H1: simulation and experimental time responses. .................................................. 59
B.2 Manual coarse tuning for subject H1: simulation and experimental amplitude spectra. ................................................. 59
B.3 Automatic fine tuning for subject H1: simulation and experimental time responses. .................................................. 60
B.4 Automatic fine tuning for subject H1: simulation and experimental amplitude spectra. ................................................. 60
B.5 Fine tuning for subject H2: simulation and experimental time responses. .................................................. 61
B.6 Fine tuning for subject H2: simulation and experimental amplitude spectra. .................................................. 61
B.7 Fine tuning for subject H3: simulation and experimental time responses. .................................................. 62
B.8 Fine tuning for subject H3: simulation and experimental amplitude spectra. .................................................. 62
B.9 Fine tuning for subject H4: simulation and experimental time responses. .................................................. 63
B.10 Fine tuning for subject H4: simulation and experimental amplitude spectra. .................................................. 63
B.11 Fine tuning for subject P1: simulation and experimental time responses. .................................................. 64
B.12 Fine tuning for subject P1: simulation and experimental amplitude spectra. .................................................. 64
B.13 Fine tuning for subject P2: simulation and experimental time responses. .................................................. 65
B.14 Fine tuning for subject P2: simulation and experimental amplitude spectra. .................................................. 65
B.15 Fine tuning for subject N1: simulation and experimental time responses. .................................................. 66
B.16 Fine tuning for subject N1: simulation and experimental amplitude spectra. ............................................. 66
B.17 Fine tuning for subject N2: simulation and experimental time responses. .................................................. 67
B.18 Fine tuning for subject N2: simulation and experimental amplitude spectra. ............................................. 67
B.19 Fine tuning for subject N3: simulation and experimental time responses. .................................................. 68
B.20 Fine tuning for subject N3: simulation and experimental amplitude spectra. ............................................. 68
B.21 Fine tuning for subject N4: simulation and experimental time responses. .................................................. 69
B.22 Fine tuning for subject N4: simulation and experimental amplitude spectra. ............................................. 69
B.23 Fine tuning for subject U1: simulation and experimental time responses. .................................................. 70
B.24 Fine tuning for subject U1: simulation and experimental amplitude spectra. ............................................. 70
B.25 Fine tuning for subject U2: simulation and experimental time responses. .................................................. 71
B.26 Fine tuning for subject U2: simulation and experimental amplitude spectra. ............................................. 71
B.27 Fine tuning for subject U3: simulation and experimental time responses. .................................................. 72
B.28 Fine tuning for subject U3: simulation and experimental amplitude spectra. ............................................. 72
B.29 Fine tuning for subject U4: simulation and experimental time responses. .................................................. 73
B.30 Fine tuning for subject U4: simulation and experimental amplitude spectra.
Abstract

An emerging trend among voice specialists is the use of quantitative protocols for the diagnosis and treatment of voice disorders. Vocal fold vibrations are directly related to voice quality. This research is devoted to providing an objective means of characterizing these vibrations. Our goal is to develop a dynamic model of vocal fold vibration, and map the parameter space of the model to a class of voice disorders; thus, furthering the assessment and diagnosis of voice disorder in clinical settings.

To this end, this dissertation introduces a new seven-mass biomechanical model for the vibration of vocal folds. The model is based on the body-cover layer concept of the vocal fold biomechanics, and segments the cover layer into three masses along the longitudinal direction of the vocal fold. This segmentation facilitates the model comparison with the motion of the vocal glottis contour derived from modern high-speed digital imaging systems. The model simulation is compared to 14 sets of experimental data from human subjects with healthy vocal folds and pathological vocal folds including nodule, polyp, and unilateral paralysis. We also propose a semi-empirical two-stage procedure for tuning the parameters so that the model response matches as closely as possible the experimental data in the time and frequency domains. The first stage involves the manual coarse tuning of parameters based on limited data to expedite the process. The second stage is an automatic (or manual) fine tuning process on a subset of the parameters tuned in the first stage based on a larger amount of data.

Once an ‘optimal’ set of model parameters has been identified, two model-based factors, quantifying the asymmetry between left and right vocal folds and anterior
and posterior segments of the vocal folds, are introduced and calculated for each of the 14 cases. The two factors form an asymmetry plane. Based on the value of the asymmetry factors for the 14 cases, the plane is subdivided into four regions corresponding to healthy vocal folds, nodule, polyp, and unilateral paralysis. This yields a clear visual aid for clinicians, correlating the model parameters to voice quality.
Chapter 1
Introduction

1.1 Motivation

Voice disorders can have significant negative effects on an individual’s quality of life. For example, they can limit one’s choice of profession or cause loss of work (temporarily or permanently), especially when the profession requires extensive use of voice (e.g., teachers, singers, actors, and news reporters). This has financial and emotional implications not only for the individual but for their family and society. It is estimated that 3% to 9% of the total population of the U.S. has a voice disorder. Of the total working population in the U.S., approximately 25% have jobs that critically require voice use with 3% requiring their voice for public safety [1]. Therefore, there is a need for objective diagnosis and treatment of voice disorders induced by vocal abuse/misuse and specific congenital, neuromuscular, or tumor-related disorders.

Vibration of the vocal folds is the primary source of voice production. Irregularity in the vocal fold vibrations may contribute to abnormal voice. An emerging trend in speech research is to correlate the characteristics of the vocal fold vibration with voice quality [41]. This is important because it can help in understanding the fundamental phonation mechanism; thereby, establishing a quantitative paradigm for diagnosing and treating voice disorders.

Presently, clinical procedures are still empirical and subjective. In particular, they rely on indirect aerodynamic and acoustic tests or direct imaging techniques (e.g., video stroboscope of the sustained phonation). These techniques can show changes in vocal fold vibratory characteristics as a disease progresses and as a re-
sult of normal aging [3, 14]. However, quantitative and objective parameters that are directly derived from the vocal fold vibration during phonation are still in their infancy. This missing information not only prevents a unified concept of what characterizes the normal vocal fold function, but also any objective determination of the effects of medical, surgical, and behavioral voice therapies. Until recently, the limitation of laryngeal imaging techniques has been a major contributor to this problem. Fortunately, recent technological improvements have led to the development of high-speed digital imaging (HSDI) systems, which permit visualization of vocal fold vibration in real time with a capture rate of 2000 to 4000 frames per second (fps). Studies employing HSDI have demonstrated that this technique gives more detailed and accurate information about the vibratory patterns of vocal folds during phonation [7, 9, 12, 20, 22].

The importance of biomechanical models of vocal fold vibration to the study of voice disorders has been recognized since the late 1950s [10, 16]. The model parameters should reflect various laryngeal and pathological configurations for physiological and clinical applications [13]. A considerable amount of work has been devoted to this subject, ranging from lumped-parameter, one- and two-dimensional models to distributed parameter, three-dimensional models [34, 35]. For instance, linear, lumped-parameter, one- and two-mass models have contributed to the understanding of vocal fold vibration, especially during normal phonation [16]. The simple one-mass model [6, 10, 11] assumes a uniform motion for the vocal fold layers and thus does not accurately model most voice vibrations. The more refined two-mass model [16, 17, 32] attempts to capture the fact that actual vocal folds have a wave-like motion from bottom to top. The two-mass model has been used to study a limited number of irregular vocal fold vibrations [16]. Multi-mass models (i.e., three or more masses) have been proposed based on the premise that the vocal fold
has two tissue layers with different mechanical characteristics – the body layer and the cover layer [33, 36]. Such multi-mass models are believed to more likely capture abnormal vibrations in the vocal fold [27]. Another type of multi-mass model was proposed in [38] by combining five two-mass models along the anterior-posterior direction of the vocal fold to account for its longitudinal flexibility.

So far, few studies have attempted to compare clinically-observed vocal fold vibrations with model simulations [25]. The limited studies can be attributed to the difficulty in obtaining and processing images of vocal fold vibrations in real time. The improved use of HSDI in a clinical setting and the development of image processing systems are now allowing further studies investigating the compatibility of vocal fold dynamic models with direct imaging of the vocal fold vibratory characteristics. This comparison will be of great value in furthering our understanding of the causes of abnormal vocal fold activities, and in predicting the effects of treatments on patients with different pathologies.

1.2 Vocal Fold Anatomy
The human vocal folds are located above the trachea and form the narrowest portion of the airway passage, named glottis[34]. The vocal folds are housed inside the larynx, a movable organ that is strengthened and upheld by cartilages and surrounding muscles. Figure 1.1 shows the sagittal view of the head and neck. From top to bottom, the nasal cavity, oral cavity, pharynx, vocal folds, and tracheal ring are the airway for respiration. The thyroid and cricoid cartilages protect the larynx and act as anchors for supporting muscles. The epiglottis is a flap that seals the entry way to the larynx during swallowing and serves as a sound resonator during phonation.
Each vocal fold is anteriorly and laterally attached to the thyroid cartilage, while its posterior is connected to the anterior angle of the arytenoid cartilage (i.e., the vocal process), as illustrated in Figure 1.2. The arytenoid cartilage, which sits on the cricoid cartilage, has the freedom to rotate and slide, allowing the adduction or abduction of the vocal folds by muscle activity. Specifically, the thyroarytenoid muscle shortens and thickens the vocal folds by contraction [34]. The cricothyroid muscle lengthens the vocal folds and is the primary pitch-control muscle. The lateral cricoarytenoid muscle and the posterior cricoarytenoid muscle control the
FIGURE 1.2. Superior view of vocal folds. Left: supporting structure; right: superior view.

vocal fold adduction and abduction, respectively. The interarytenoid muscle can help narrow the posterior glottis.

Histologically, the vocal folds are composed of three layered structures [13] as shown in Figure 1.3: the epithelium, the lamina propria, and the thyroarytenoid muscle. The lamina propria, which is made up of nonmuscular connective tissues, is subdivided into three layers: superficial, intermediate, and deep layers. The superficial layer is pliable with a loose fibrous structure that allows for large elongations. The intermediate layer is composed of elastin fibers and collagen fibers along the anterior and posterior directions [34]. The deep layer has more collagen fibers, which provides more strength and rigidity. Due to the flexible nature of the superficial and intermediate layers during phonation, these two layers together with
epithelium are referred to as the vocal fold *cover*, while the more rigid structures, viz., the deep layer and the thyroarytenoid muscle, form the vocal fold *body* [15].

### 1.3 Voice Production

The vibration of the vocal folds is the primary source of sound production [34]. When the subglottal pressure overcomes the so-called phonation threshold pressure due to an increase in the lung pressure, the vocal folds begin to oscillate in a self-sustained manner. A typical cycle can be described as follows. The pliable nature of the vocal fold cover allows the subglottal pressure to first push the lower part of the vocal folds laterally. The upper part of vocal folds is then dragged into motion by the lower part. During this outward movement, the glottis is in a convergent shape which allows a positive pressure to continuously act upon the lower portion of vocal folds according to Bernoulli’s law. As the vocal folds reach their maximum displacement, the elastic force of the vocal fold becomes strong enough and pushes the vocal folds inward. Since the lower part is leading the motion, the glottis now has a divergent shape, which makes the pressure on the lower part of vocal folds nearly zero. The lack of pressure causes the right and left vocal folds to collide.
with each other, closing the glottis. With the glottis closed, the pressure builds up again and another cycle begins. It is this leading motion of the lower portion of vocal folds, which is referred to as the “ribbon mode” or a “wave-like” movement [34], that allows the subglottal pressure to transmit energy to the vocal folds and sustain the vibration.

The vibrating vocal folds cause the air density near the outlet of the vocal folds to cyclically increase and decrease (i.e., air condensation and rarefication). This disturbance of the air density propagates back to the vocal trachea and forward to the vocal tract. The vocal tract acts as a sound filter, amplifying selected frequencies and suppressing others. Changes in the length and shape of the vocal tract (e.g., tongue position and mouth shape) modulate the filter, modifying the frequencies amplified/suppressed[34].

1.4 Voice Disorders

In this research, we will focus on characterizing the dynamics of three voice disorders: vocal fold nodules, vocal fold polyp, and unilateral vocal fold paralysis.

Nodules are usually a benign lesion caused by the collision of the vocal folds during phonation. As such, vocal nodules are symmetric and located in the middle part of the vocal edges. Excessive collision of the vocal folds is believed to lead to mechanical stresses [5, 34]. In the initial stage, when nodules first appear, they are soft and pliable. If the abusive use of the vocal folds is prolonged, the edema formed in early stages will undergo fibrosis and the nodules will harden and cause hypertrophy of the epithelium. Acoustically, patients with nodules exhibit some degree of dysphonia, jitter, and shimmer [5].

Polyp is usually caused by a localized irritation, such as smoke, chemicals, or internal rupture of a blood vessel [34]. They are mainly unilateral. Compared
with nodules, polyps are typically larger, hemorrhagic, fibrotic, and inflammatory. Acoustic signs of polyps and nodules are similar, however. Asymmetric motion of the vocal folds can be observed by stroboscopy due to its unilateral occurrence [5].

Vocal fold paralysis is immobility of one or both vocal folds mostly caused by lesions on the vagus nerve of the larynx (superior laryngeal nerve, recurrent laryngeal nerve, or both). Clinical studies have found that unilateral paralysis is more common than bilateral paralysis, accounting for 79.8% of paralysis cases [5]. Acoustically, patients with unilateral paralysis usually have breathy and hoarse voice. Due to incomplete glottal closure, subglottal pressure and air flow are much higher than normal [23].

1.5 Vocal Fold Examination

Vocal folds are usually examined via acoustical or imaging systems. With the acoustical system, a variety of vocal fold characteristics can be obtained such as fundamental frequency, phonation range, vocal intensity, and acoustic spectrum [5]. The subject usually is instructed to produce simple vowels such as /ah/ or /ee/, or a long reading or conversation. These acoustical signals are recorded by microphone and simple digital data acquisition devices. Sound analysis software is then applied to further manipulate the recorded signals.

Videostroboscopy uses a strobe to slow down the rapid vibrating vocal fold image. However, this method has a capture rate of only 35 fps, which is not enough to accurately examine the cycle of vibration. On the other hand, HSDI produces 2000 to 4000 fps and gives detailed phonation information. This technique first emerged at the Bell Telephone Laboratories in 1937 [5], but only became affordable for clinical application with the advent of high-speed digital cameras and digital data acquisition techniques. Figure 1.4 illustrates a HSDI system with a rigid endo-
scope from KayPENTAX (http://kaypentax.com). In this model, the endoscope is connected to a constant Xenon light source that illuminates the larynx area. A high-speed digital camera, attached to the tip of endoscope, is capable of capturing 2000 fps with a 120 × 256 pixel resolution for a duration of 2.2 seconds. This means that, if the average fundamental frequency of the vocal fold vibration is 200 Hz, approximately 10 frames will be captured during each cycle. During testing, the endoscope is inserted through the oral cavity of the subject until it reaches the larynx. The subject is then instructed to pronounce a sustained high-pitch /ee/ vowel, which moves the tongue upward and forward. This opens the larynx and lengthens the vocal folds, allowing the whole superior view of the vocal folds to be better exposed to the camera. In addition to a video of the vocal fold vibration, HSDI systems can provide a synchronized acoustical signal and electroglottography. A typical vibration cycle captured by the HSDI system is illustrated in Figure 1.5.
1.6 Scope of Work

This research seeks to characterize the dynamic properties of vocal fold vibrations of normal and disordered voices. The outcome of this research is the mapping of the parameter space of the vocal fold model to a class of voice disorders (viz., nodule, polyp, and unilateral paralysis); thus, furthering the assessment and diagnosis of voice disorder in clinical settings. A block diagram of the proposed dynamic characterization scheme is shown in Figure 1.6. Clinical data of the actual vocal fold vibration during phonation is obtained via a HSDI system, yielding an output signal (e.g., position of points along the vocal fold edge). A dynamic model for the vocal fold vibrations is defined, parameterized by the parameter vector $\theta$. A semi-empirical parameter tuning method for the parameter vector $\theta$ is introduced based on the error between the actual (experimental) and model outputs. The set of parameters $\{\theta_i\}$, where $\theta_i$ represents the parameter vector for subject (voice
disorder) i, will then be used to classify regions of the parameter space vis-à-vis the voice pathology.

This work was conducted in collaboration with Dr. Melda Kunduk of the LSU Department of Communication Sciences and Disorders, graduate student Jing Chen of the LSU Department of Electrical and Computer Engineering, and Dr. Bahadir Gunturk of the LSU Department of Electrical and Computer Engineering. A preliminary version of this research appeared in [37].
Chapter 2
Biomechanical Models of Vocal Folds

In this chapter, we provide a literature review of discrete (i.e., lumped-parameter) vocal fold vibration modeling.

2.1 One-Mass Model

The first discrete model for the vocal fold dynamics was the simple one-mass model [6, 10, 11]. In this model, each vocal fold is represented by a mass attached to a linear spring and a linear viscous damper as shown in Figure 2.1. In the case where the left and right vocal folds are identical (i.e., symmetric vocal folds), the dynamics are characterized by the simple second-order equation

\[ m\ddot{x} + b\dot{x} + kx = F, \]

(2.1)

where \( x \) is the lateral displacement of the vocal fold, \( m \) is the total mass of the vocal fold, \( b \) represents the viscosity of the vocal fold tissue, \( k \) represents the elasticity of the vocal fold tissue, and \( F \) is the laterally-applied aerodynamic forcing function. If the length of glottis (in the anterior-posterior direction) is \( L \) and the depth of the vocal fold (in the flow direction) is \( d \), the forcing function \( F \) is determined by the average of the inlet pressure \( P_1 \) and outlet pressure \( P_2 \) of the glottal orifice [11]:

\[ F = \frac{Ld}{2} (P_1 + P_2). \]

(2.2)

The pressures \( P_1 \) and \( P_2 \) are given by the following empirical formulae [11]

\[ P_1 = P_s - 0.685 \rho \frac{u_g^2}{a^2} \]

(2.3)

\[ P_2 = -0.25 \rho \frac{u_g^2}{a^2} \]

(2.4)
where $P_s$ is the subglottal pressure, $u_g$ is the volume flow rate in the glottis, $\rho$ is the air density, $a = 2(x + x_0)L$ is the glottal area, and $x_0$ is the mass rest position. The volume flow rate is a function of $P_s$ and $a$, i.e., $u_g = u_g(P_s, x)$, through the coupling of (2.1) to the equations for the acoustic circuit of voice production; see [11] for details.

Since the simple one-mass model assumes a uniform, one-dimensional motion for the vocal fold layers, it does not accurately model most voice vibrations. Specifically, it is incapable of explaining how flow energy is transferred to the vocal fold tissue to sustain the oscillation [34].

### 2.2 Two-Mass Model

The two-mass model shown in Figure 2.2 is the most widely used biomechanical model of vocal fold vibrations. The fundamental reason for its popularity is that it captures the phase difference in the displacements of the lower and upper edges of the vocal fold observed in real vocal fold oscillations. That is, the two-mass model
captures the fact that actual vocal folds have a wave-like motion from bottom to top. This is illustrated in Figure 2.3, where we can see that the lower edge of the vocal fold always leads the motion.

In the two-mass model, the masses do not strictly reflect the anatomical or physiological structure of vocal folds, and are not directly related to vocal tissues [17]. The model assumes each vocal fold is represented by a pair of damped mechanical oscillators coupled by a spring [17]. The glottis is approximated by rectangular cross sections [2]. The two-mass model usually also includes the deformation of the vocal folds during collision of the left/right folds. The two-mass model was introduced in 1972 by [16] and later simplified in [32]. In original version of the model, the mechanical and aerodynamical interactions were expressed as complex coupled ordinary differential equations with many parameters [32]. The simplifying assumptions are the following [7, 18, 32]:

- The cubic nonlinearity that models the elastic property of the vocal fold tissue is neglected.
- The influence of tract and subglottal resonances is neglected.
- Subglottal and supraglottal resonances are negligible.
- The pressure drop at the inlet due to vena contracta is neglected.
- Viscous losses inside the glottis are neglected.
- The pressure force of the Bernoulli flow affects up to the narrowest part of the glottis.
- The vocal folds are symmetric.
Based on above assumptions, the two-mass model for each vocal fold is given by

the following equations of motion [32]

\[
\begin{align*}
    m_1 \ddot{x}_1 + b_1 \dot{x}_1 + k_1 x_1 + k_c (x_1 - x_2) + \Theta(a_1) c_1(x_1 + x_{01}) &= F_1(x_1, x_2) \quad (2.5a) \\
    m_2 \ddot{x}_2 + b_2 \dot{x}_2 + k_2 x_2 + k_c (x_2 - x_1) + \Theta(a_2) c_2(x_2 + x_{02}) &= F_2(x_1, x_2), (2.5b)
\end{align*}
\]

where \( \Theta \) is a “switching” function that is turned on during collision of the left and
right vocal folds [24], and is given by

\[
\Theta(a_i) = \begin{cases} 
1, & \text{if } a_i \leq 0 \\
0, & \text{if } a_i > 0.
\end{cases} \quad (2.6)
\]

The parameters and variables in (2.5) and (2.6) are defined in the table below.
FIGURE 2.3. A cycle of the vocal fold vibration showing the wave-like motion from bottom to top. Left column shows the opening of the glottis and right column shows the closing. (http://voiceproblem.org/glossary/images_anatomy1.asp)

\[ x_i \quad \text{Lateral displacement of mass } m_i \text{ from midline} \]
\[ x_{0i} \quad \text{Rest position of } m_i \text{ from midline} \]
\[ a_i = 2l(x_i + x_{0i}) \quad \text{Cross sectional area of } m_i \]
\[ a_{0i} = 2lx_{0i} \quad \text{Cross sectional area of } m_i \text{ at rest position} \]
\[ m_i \quad \text{Mass } i \]
\[ k_i \quad \text{Stiffness constant} \]
\[ k_c \quad \text{Coupling stiffness constant} \]
\[ c_i \quad \text{Collision stiffness constant} \]
\[ b_i \quad \text{Damping constant} \]
\[ L \quad \text{Anterior-posterior length of the glottis} \]
\[ d_i \quad \text{Depth of } m_i \]
The forcing function $F_i(x_1, x_2)$ in (2.5) is given by

$$F_i = Ld_i P_i$$  \hspace{1cm} (2.7)

where $P_i$ is the glottal pressure acting on $m_i$. Since $P_2$ is close to the outlet pressure of the glottal orifice, which is approximately zero, we typically set $P_2 = 0$ so that $F_2 = 0$. To obtain $F_1$, we consider the following four cases.

**Case 1.** Glottal channel is convergent, i.e., $a_1 > a_2$.

In this case, the airflow at upper mass region is given by the Bernoulli equation

$$P_s = P_1 + \frac{\rho u_g^2}{2a_1^2} = P_2 + \frac{\rho u_g^2}{2a_2^2}.$$  \hspace{1cm} (2.8)

Since $P_2 = 0$, $P_1$ can be expressed as

$$P_1 = P_s \left[1 - \left(\frac{a_2}{a_1}\right)^2\right].$$  \hspace{1cm} (2.9)

**Case 2.** Glottal channel is divergent, i.e., $0 < a_1 \leq a_2$.

Here, the airflow detaches from the surface of the vocal folds and forms a jet, which keeps the pressure in the glottis closer to zero. Thus,

$$P_1 = 0.$$  \hspace{1cm} (2.10)

**Case 3.** Upper masses $m_2$ collide, while lower masses $m_1$ remain apart, i.e., $a_1 > a_2$ and $a_2 \leq 0$.

Due to the absence of flow through the glottal orifice, we have from (2.8) that

$$P_1 = P_s.$$  \hspace{1cm} (2.11)

**Case 4.** Lower part masses collide, i.e., $a_1 \leq 0$.

This is similar to case 2, hence

$$P_1 = 0.$$  \hspace{1cm} (2.12)
The above cases can be compactly written as follows

$$P_1 = P_s \left[ 1 - \Theta(-a_{min}) \left( \frac{a_{min}}{a_1} \right)^2 \right] \Theta(-a_1) \quad (2.13)$$

where $a_{min} = \min(a_1, a_2)$ and $\Theta$ was defined in (2.6). Further, the glottal volume flow rate for all cases is given by the single equation

$$u_g = \sqrt{\frac{2P_s}{\rho} a_{min} \Theta(a_{min})} \quad (2.14)$$

Although the two-mass model is effective in simulating the vibration of normal vocal folds, it can only reproduce a limited class of irregular vocal fold vibrations [16]. Thus, further refinements to the biomechanical model are needed to facilitate medical diagnosis and treatment of voice disorders.

### 2.3 Multi-Mass Models

The term “multi-mass model” is used in reference to vocal fold models that contain three or more masses per vocal fold. One such model was recently presented in [40] by coupling 3 two-mass models via springs along the length of the vocal fold to create a two-dimensional model. Specifically, two-mass models were located at 25%, 50%, and 75% of the glottal length to represent the dorsal, medial, ventral motions of the vocal fold [40]. This model accounts for the fact that the vocal fold is flexible along the anterior-posterior direction, similar to a vibrating string.

Although the terms *vocal cord* and *vocal fold* are often used interchangeably, the term *vocal fold* is more appropriate since the structure of the vocal tissue is cover-body layered as explained in Section 1.2. Further, when vibrating, the cover layer is usually delayed and out of phase with the body layer. That is, the vocal fold is not as simple as a vibrating string [36]. Multi-mass models have been proposed based on the premise that the two tissue layers have different mechanical characteristics. Such multi-mass models are more likely to reproduce abnormal vibrations in the
vocal fold [27] in comparison to the two-mass model. The three-mass model shown in Figure 2.4 was introduced by [33] in 1995. In this model, masses $m_1$ and $m_2$ represent the wave-like motion of the cover layer (as in the two-mass model) while $m_3$ along with $k_3$ and $r_3$ represent the dynamics of the body layer. Note that this model is one dimensional like the standard two-mass model. The equations of motion for the three-mass model of one vocal fold, under the same assumptions listed in Section 2.2, are the following

$$m_1 \ddot{x}_1 + r_1 (\dot{x}_1 - \dot{x}_3) + k_1 (x_1 - x_3) + k_c (x_1 - x_2) + \Theta(a_1) c_1 (x_1 + x_{01}) = F_1(x_1, x_2)$$

$$m_2 \ddot{x}_2 + r_2 (\dot{x}_2 - \dot{x}_3) + k_2 (x_2 - x_3) + k_c (x_2 - x_1) + \Theta(a_2) c_2 (x_2 + x_{02}) = F_2(x_1, x_2)$$

$$m_3 \ddot{x}_3 + r_1 (\dot{x}_3 - \dot{x}_1) + r_2 (\dot{x}_3 - \dot{x}_2) + r_3 \dot{x}_3 + k_1 (x_3 - x_1) + k_2 (x_3 - x_2) + k_3 x_3 = 0,$$

where $F_i$ and $\Theta(a_i)$, $i = 1, 2$ are defined as in Section 2.2.

### 2.4 Asymmetric Models

Some voice disorders induce asymmetry in the biomechanical properties of the left and right vocal folds. As a result, the parameters in the model (2.5), for example, are different for each vocal fold. Several researchers have expanded the two-mass...
model to study vibration with asymmetric variations in tension, resting glottal gap, and subglottal pressure [17]. One asymmetric model is based on the asymmetry factor $Q$ defined as [8, 25]

$$k_{ir} = Qk_{il} \quad \text{and} \quad m_{ir} = \frac{m_{il}}{Q}, \quad i = 1, 2. \quad (2.16)$$

As the fundamental frequencies of the two-mass model oscillations can be approximated by the simple mass-spring oscillator, it follows from (2.16) that

$$f_r = Qf_l \quad (2.17)$$

where $f_r, f_l$ are the vibration frequency of the right and left vocal folds, respectively. For normal voices, it has been experimentally verified that $0.95 \leq Q \leq 1.05$ for females and $0.91 \leq Q \leq 1.10$ for males [8]. That is, symmetry in the left/right vocal fold vibrations ($Q \approx 1$) is an important indicator of normal voice. Another related approach for incorporating asymmetry into the two-mass model is to use two asymmetry factors $Q_r$ and $Q_l$ defined as [7, 32]

$$k_{i\alpha} = Q_{\alpha}k_{i\alpha 0} \quad \text{and} \quad m_{i\alpha} = \frac{m_{i\alpha 0}}{Q_{\alpha}}, \quad i = 1, 2; \quad \alpha = r, l \quad (2.18)$$

where the subscript 0 denotes the standard parameter values. These so-called standard values are taken from the work in [16], which established relevant ranges for the parameters of the two-mass model for normal voice. The fundamental frequencies are then approximated as

$$f_{\alpha} = Q_{\alpha}f_{\alpha 0}, \quad \alpha = r, l \quad (2.19)$$

where $f_{\alpha 0}$ is the “standard” vibration frequency of the $\alpha$ vocal fold [7, 32]. A time-dependent version of (2.18) was discussed in [40].

A factor for quantifying the asymmetry between the anterior and posterior segments of a vocal fold, similar to (2.19), was introduced in [27]. In [39], factors were proposed comprising both left/right and anterior/posterior asymmetries.
2.5 Estimation Approaches for Model Parameters

An important topic of research in vocal fold modeling is identifying the model parameters and correlating them with voice disorders. Here, we review previous vocal fold modeling work where parameter estimation-like methods were used to automatically identify the model parameters off-line.

In [7], a frequency-domain parameter estimation method was proposed to identify parameters of the two-mass model based on HSDI recordings of the vocal fold vibrations. The time evolution of the distance between the medial position of the left and right vocal folds were extracted from the model and HSDI recordings, yielding theoretical and experimental curves, respectively. A nonconvex objective function was defined as the difference between the dominant Fourier coefficients of the theoretical and experimental curves, and the Nelder-Mead algorithm was used for the minimization procedure. In [31], a Genetic Algorithm was used to identify the subglottal pressure, masses, spring constants, and mass rest positions of the multi-mass model proposed in [40]. The objective function was defined in the time domain as a combination of two criteria that quantify the agreement between the experimental and theoretical curves: the difference in glottal areas and the difference in distance of the dorsal, medial, and ventral vocal edges from the glottal axis. In [39], the dynamics of the multi-mass model of [31] with time-varying parameters were matched to vocal fold vibrations at the dorsal, medial, and ventral positions using the discrete wavelet transform and Powell’s direct set method.
Chapter 3
Proposed Biomechanical Model

In this chapter, we introduce a new, seven-mass biomechanical model for the vibration of vocal folds. The model is based on the body-cover layer concept of the vocal fold biomechanics, and segments the cover layer into three masses.

3.1 Basic Nomenclature

Before presenting the differential equations that govern the dynamics of the proposed multi-mass model, we introduce the following nomenclature:

- Superscript $\alpha = r, l$ Right and left vocal fold, respectively
- Subscript $\beta = u, l$ Upper and lower cover layer, respectively
- Subscript $i = 1, 2, 3$ Ventral, medial, and dorsal segment, respectively
- $m^\alpha_B$ Mass of body layer
- $m^\alpha_{\beta i}$ Mass of cover layer
- $k^\alpha_{\beta i}$ Stiffness coefficient between body masses and cover mass
- $k^\alpha_B$ Stiffness coefficient between body mass and cartilage
- $b^\alpha_{\beta i}$ Damping coefficient between body masses and cover mass
- $b^\alpha_B$ Damping coefficient between body mass and cartilage
- $\kappa^\alpha_{s i}$ Stiffness coefficient between upper masses and lower cover masses
- $\kappa^\alpha_{\beta j}$ Stiffness coefficient between cover masses or cover masses and cartilage, $j = 1, 2, 3, 4$
- $c_{\beta i}$ Stiffness coefficient between right and left masses due to collision
- $d_i$ Depth of the $i$th lower mass (in the direction of airflow)
- $L_i$ Anterior-posterior length of the $i$th lower mass
\[ x_B^\alpha \] Displacement of mass \( m_B^\alpha \) relative to rest position

\[ x_{\beta_i}^\alpha \] Displacement of mass \( m_{\beta_i}^\alpha \) relative to rest position

\[ d_{\beta_i}^\alpha \] Rest position of mass \( m_{\beta_i}^\alpha \) from glottal axis

\[ F_{\beta_i} \] Force on mass \( m_{\beta_i}^\alpha \) due to airflow through glottis

\[ P_i \] Pressure inside the glottis on lower mass \( m_i^\alpha \)

\[ P_s \] Subglottal pressure

\[ a_\beta \] Glottis area

\[ a_{\beta_i} \] Portion of glottis area between \( m_{\beta_i}^r \) and \( m_{\beta_i}^l \)

\[ a_{\beta_i0} \] Portion of glottis area between \( m_{\beta_i}^r \) and \( m_{\beta_i}^l \) at rest position

### 3.2 Model Description

We developed a two-dimensional, multi-mass model which can be viewed as the combination of the three-mass model and the multi-mass model of [40]. Specifically, our model uses 3 two-mass models at 25%, 50%, and 75% of the glottal length connected to a single, “third” mass. The side and top views of the proposed multi-mass model are shown in Figures 3.1 and 3.2, respectively. A 3-D view of the model is shown in Figure 3.3. Note that this model incorporates the three main characteristics of the vocal fold dynamics:

- The wave-like motion of the cover layer from bottom to top.
- The different biomechanical properties of the body and cover tissues.
- The flexibility of the cover layer in the anterior-posterior direction.

The use of only 3 two-mass models along the anterior-posterior direction is motivated by the trade-off between model fidelity and number of model parameters in need of identification.
The equations of motion for the lower cover masses of the vocal folds are

\[
m_{l1} \ddot{x}_{l1} + b_{l1}^r (\dot{x}_{l1} - \dot{x}_{B}^r) + k_{l1}^r (x_{l1} - x_{B}^r) + \kappa_{l1}^r (x_{l1} - x_{l2}^r) + \kappa_{s1}^r (x_{l1} - x_{u1}^r) + \Theta(-x_{l1}^r - d_{l1}^r - x_{l1}^l - d_{l1}^l) c_{l1} (x_{l1}^l + d_{l1}^l + x_{l1}^l + d_{l1}^l) = F_{l1}
\]

\[
m_{l2} \ddot{x}_{l2} + b_{l2}^r (\dot{x}_{l2} - \dot{x}_{B}^r) + k_{l2}^r (x_{l2} - x_{B}^r) + \kappa_{l1}^r (x_{l2} - x_{l1}^r) + \kappa_{l2}^r (x_{l2} - x_{l3}^r) + \kappa_{s2}^r (x_{l2} - x_{u2}^r) + \Theta(-x_{l2}^r - d_{l2}^r - x_{l2}^l - d_{l2}^l) c_{l2} (x_{l2}^l + d_{l2}^l + x_{l2}^l + d_{l2}^l) = F_{l2}
\]

\[
m_{l3} \ddot{x}_{l3} + b_{l3}^r (\dot{x}_{l3} - \dot{x}_{B}^r) + k_{l3}^r (x_{l3} - x_{B}^r) + \kappa_{l4}^r (x_{l3} - x_{l4}^r) + \kappa_{i4}^r (x_{l3} - x_{l4}^l) + \kappa_{s3}^r (x_{l3} - x_{u3}^r) + \Theta(-x_{l3}^r - d_{l3}^r - x_{l3}^l - d_{l3}^l) c_{l3} (x_{l3}^l + d_{l3}^l + x_{l3}^l + d_{l3}^l) = F_{l3}
\]

where

\[
\Theta(\eta) = \begin{cases} 
1, & \eta > 0 \\
0, & \eta \leq 0,
\end{cases}
\]

\[
F_i = L_i d_i P_i,
\]
FIGURE 3.2. Top view of the proposed multi-mass model. Only upper masses of cover layer are shown.

\[ P_i = \begin{cases} 
P_s \left[ 1 - \left( \frac{a_u}{a_l} \right)^2 \right], & \text{if } a_l > a_u > 0 \\
0, & \text{if } 0 < a_l < a_u \text{ or } a_l < 0 \\
P_s, & \text{if } a_u < 0 \text{ and } a_l > 0, 
\end{cases} \quad (3.6) \]

\[ a_\beta = \sum_{i=1,2,3} a_{\beta i}, \quad (3.7) \]

and

\[ a_{\beta i} = \begin{cases} 
L_i(x^r_{\beta i} + x^l_{\beta i}) + a_{\beta i0}, & \text{if there is no collision} \\
0, & \text{if collision occurs.} 
\end{cases} \quad (3.8) \]

The switching function \( \Theta \) in (3.4) accounts for the possible collision of the left and right masses during the glottis closing. Due to the lateral compliance of the vocal fold cover tissue, the collision is modeled as a linear spring. The equations for the three upper cover masses are similar to (3.1)-(3.3) with the exception that \( F_{ui} = 0, \)
The equations of motion for the body masses are

\[ m_B^\alpha \ddot{x}_B^\alpha + b_B^\alpha \dot{x}_B^\alpha + k_B^\alpha x_B^\alpha + \sum_{\beta=l,u, i=1,2,3} b_{\beta i}^\alpha (\dot{x}_B^\alpha - \dot{x}_{\beta i}^\alpha) + k_{\beta i}^\alpha (x_B^\alpha - x_{\beta i}^\alpha) = 0 \quad (3.9) \]

Note that the above model is nonlinear due to the switching function \( \Theta \) and the input nonlinearity given by (3.6) (i.e., \( F_{li} \) is a nonlinear function of \( x_{\beta i}^\alpha \)).

For illustration purposes, the above mathematical model was simulated in MATLAB for symmetric (healthy) vocal folds having the following nominal parameters
\[ m_B^\alpha = 0.05 \text{ g}, \quad m_{\beta i}^\alpha = 0.0033 \text{ g}, \]
\[ k_{ii}^\alpha = 0.0017 \text{ g/ms}^2, \quad k_{ui}^\alpha = 0.0012 \text{ g/ms}^2, \quad k_B^\alpha = 0.1 \text{ g/ms}^2, \]
\[ b_{li}^\alpha = 1.9 \times 10^{-3} \text{ g/ms}, \quad b_{ui}^\alpha = 1.6 \times 10^{-3} \text{ g/ms}, \quad b_B^\alpha = 0.0283 \text{ g/ms}, \]
\[ \kappa_{si}^\alpha = 0.00067 \text{ g/ms}^2, \quad \kappa_{li}^\alpha = 0.0083 \text{ g/ms}^2, \]
\[ \kappa_{ui}^\alpha = 0.0058 \text{ g/ms}^2, \quad c_{\beta i} = 3k_{\beta i}^\alpha, \]
\[ L_i = 0.4 \text{ cm}, \quad w_i = 0.15 \text{ cm}, \]
\[ d_{l1}^\alpha = 0.0045 \text{ cm}, \quad d_{l2}^\alpha = 0.018 \text{ cm}, \quad d_{l3}^\alpha = 0.032 \text{ cm}, \]
\[ d_{u1}^\alpha = 0.00447 \text{ cm}, \quad d_{u2}^\alpha = 0.0179 \text{ cm}, \quad d_{u3}^\alpha = 0.0313 \text{ cm}, \]
\[ P_s = 0.008 \text{ g/(cm-ms}^2). \]

All the initial conditions were set to zero, except for \( x_{\beta i}^\alpha = 0.1 \text{ cm}. \) The simulation results for the displacements of the seven masses are shown in Figure 3.4. Due to the symmetry of the simulated vocal folds, only the vibrations of one side are shown. Notice how the lower masses always lead the motion relative to the upper masses, as expected.
FIGURE 3.4. Simulation results of symmetric vocal folds using the proposed multi-mass model.
Chapter 4
Tuning of Model Parameters

In order to validate the biomechanical model given by (3.1)-(3.9), we will evaluate its ability to reproduce actual vibrations of healthy and pathological (nodule, polyp, and unilateral paralysis) vocal folds by comparing the model response to experimental data obtained from a HSDI system. The model parameters (viz., masses, stiffnesses, damping coefficients, mass rest positions, and subglottal pressure) directly influence the model response. Therefore, in this chapter, we will introduce semi-empirical procedures for tuning the parameters so that the model response matches as closely as possible experimental data from given vocal folds. Specifically, we seek to match the displacement in time and fast Fourier transform (FFT) data of the ventral, medial, and dorsal segments of the vocal fold edges.

We present a two-stage parameter tuning approach. The first stage involves the manual coarse tuning of parameters based on a data set consisting of 90 ms of simulation time, 64 ms of experimental time, and a 128-point FFT for both the simulation and experimental data. These values were used to expedite the manual course-tuning process. This tuning stage is initialized with a nominal parameter set for healthy vocal folds, viz., the parameter set in (3.10). At the end of this stage, we obtain a model parameter set whose responses are relatively close to their experimental counterparts. Since the 128-point FFT is not very accurate, we then run an automatic fine-tuning process on a subset of the parameters tuned in stage one using 280 ms of simulation time, 256 ms of experimental time, and a 512-point FFT.
Due to the inability of the HSDI system to distinguish the lower and upper edges of the vocal fold (see Section 5.1), the model simulation needs to mimic this fact to enable a proper comparison with the experimental data. To this end, we obtain the contour of the vocal folds in the simulation by setting the left and right vocal fold edges to

$$
x_i^\alpha := \begin{cases} 
\min(x_{li}^\alpha + d_{li}^\alpha, x_{ui}^\alpha + d_{ui}^\alpha), & \text{if section } i \text{ of glottis is open} \\
\text{Position of the collision,} & \text{if collision occurs in section } i.
\end{cases}
$$

(4.1)

Thus, we use the vocal fold segment variables $x_i^\alpha$ to compare the model simulation with the experimental data.

### 4.1 Manual Coarse-Tuning Procedure

To reduce the number of parameters to be tuned, the damping coefficients $b_{\beta i}^\alpha$ and $b_B^\alpha$ in the model were set to [33]

$$
b_{\beta i}^\alpha = 2\zeta_1 \sqrt{m_{\beta i}^\alpha k_{\beta i}^\alpha}
$$

(4.2)

and

$$
b_B^\alpha = 2\zeta_2 \sqrt{m_B^\alpha k_B^\alpha}
$$

(4.3)

where the damping ratio $\zeta_1$ is set to a value in the interval $[0.3, 0.4]$ if there is no collision and to a value in the interval $[0.6, 0.8]$ if collision occurs, and

$$
\zeta_2 = 0.4.
$$

(4.4)

The manual tuning of the model parameters begins with an initial, standard parameter set; e.g., the parameter set in (3.10). For each change of parameters, the simulation outputs the FFT and time domain behavior of $x_i^\alpha$. The tuning of each parameter from its initial value is guided by the following physical observations and insights.
• The fundamental frequency of the steady-state vibrations of (3.1)-(3.9) is partially dependent on the fundamental frequency of the Bernoulli pressure $P_i$. However, the Bernoulli pressure depends on the relative displacement of upper and lower masses (see (3.6)-(3.8)), complicating the determination of its fundamental frequency. Fortunately, we observed after running several simulations with different parameter values that the fundamental frequency of the steady-state vibrations is very close to the one for the unforced system, which is obtained by setting $F_{li} = 0$ in (3.1)-(3.3). This observation facilitates the parameter tuning since we can assume that the vibration frequencies are proportional to $\sqrt{k/m}$ where $k$ is the spring constant and $m$ is the mass.

• The amplitude of the vocal fold vibration is mostly sensitive to the tension of the vocal folds close to the edge and to the subglottal pressure. Increasing the subglottal pressure increases the amplitude, while increasing the tension decreases the amplitude. Since only lateral motions are considered in the proposed model, the vocal fold tension is represented by the spring constants $\kappa_{\beta i}^\alpha$.

• Asymmetric configuration of masses causes phase differences between left and right displacement. The side with lighter mass will lead the side with heavier mass.

• The collision of the vocal folds is mostly affected by the collision springs $c_{\beta i}$, the stiffness coupling the upper and lower masses $\kappa_{si}^\alpha$, and the rest position of the masses $d_{\beta i}^\alpha$.

Based on the above facts, the manual coarse tuning of the model parameters was conducted according to the following procedure:
1. Tune $m_{\beta_i}^\alpha$, $k_{\beta_i}^\alpha$, and $\kappa_{\beta_i}^\alpha$ until the frequency of the first harmonic in the FFT of $x_i^\alpha$ roughly (visually) matches the corresponding experimental value.

2. Tune $k_{\beta_i}^\alpha$, $\kappa_{\beta_i}^\alpha$ and $P_s$ until the amplitude of the first harmonic in the FFT of $x_i^\alpha$ roughly matches the corresponding experimental value. If the frequencies of the first harmonic shift as a result, go back to step 1.

3. Tune $\kappa_{s_i}^\alpha$, $c_{\beta_i}$, and $d_{\beta_i}^\alpha$ until the duration of collisions and the vibration amplitudes of $x_i^\alpha$ in time roughly match the corresponding experimental value. If the frequencies of the first harmonic shift as a result, go back to step 1.

4. Tune $m_{\beta_i}^\alpha$ until phase difference between $x_i^r$ and $x_i^l$ in time roughly matches the corresponding experimental value. If the frequencies of the first harmonic shift as a result, go back to step 1.

4.2 Automatic Fine-Tuning Procedure

In this tuning stage, we introduce some variables to make the comparison between the model and experimental data more objective and to automate the tuning process. To this end, we define

$$\tilde{f} = \frac{\sum_\alpha \sum_i \left[ f_{0i}^\alpha \right]_{\text{exp}} - \left[ f_{0i}^\alpha \right]_{\text{model}}}{6}$$

(4.5)

where $f_{0i}^\alpha$ is the fundamental frequency (first harmonic) of the FFT of $x_i^\alpha$. Let $\bar{x}_i^\alpha(j)$ and $\underline{x}_i^\alpha(j)$ be the maximum and minimum displacement of $x_i^\alpha$ in cycle $j$, respectively, as shown in Figure 4.1. The average peak-to-peak displacement amplitude is defined as

$$h_i^\alpha = \frac{1}{N_c} \sum_{j=1}^{N_c} \left[ \bar{x}_i^\alpha(j) - \underline{x}_i^\alpha(j) \right]$$

(4.6)

where $N_c$ is number of cycles in the data set. The error in $h_i^\alpha$ is then given by

$$\tilde{h}_i^\alpha = [h_i^\alpha]_{\text{exp}} - [h_i^\alpha]_{\text{model}}.$$
For those ventral and medial displacements which have large collision duration, the second and higher order harmonic components are relatively large, \( h_i^\alpha \) shows the effect of summation of all major harmonic components. Small \( \tilde{h}_i^\alpha \) indicate good resemblance between experiment and simulation data. This avoids complicate comparison of all major harmonic in FFT domain in both amplitude and phase. Also usage of \( \tilde{h}_i^\alpha \) as comparison variables has some sense of compensation, when difference of collision duration between simulation and experiment are large.

Finally, we define

\[
\tilde{A}_i^\alpha = [A_{0i}]_{\exp} - [A_{0i}]_{\text{model}}
\]

(4.8)

where \( A_{0i}^\alpha \) denotes the amplitude of the fundamental frequency (first harmonic) of the FFT of \( x_i^\alpha \).

In this stage, we only tune the parameters \( m_{\beta i}^\alpha, k_{\beta i}^\alpha, \) and \( \kappa_{\beta i}^\alpha \). The procedure begins with the parameter set that resulted from the manual coarse tuning. The
algorithm for the fine-tuning process is as follows. The flowchart is shown in Figure 4.2.

1. Multiply \( m_{\beta i}^\alpha \) and divide \( k_{\beta i}^\alpha \) and \( \kappa_{\beta i}^\alpha \) by a factor \( \rho \) until \( |\tilde{f}| \leq \varepsilon_f \) where \( \varepsilon_f \) is a sufficiently small, user-defined tolerance.

2. Multiply \( m_{\beta i}^\alpha \), \( k_{\beta i}^\alpha \), and \( \kappa_{\beta i}^\alpha \) by a factor \( \sigma^\alpha \) until \( |\tilde{h}_{i}^\alpha| \leq \varepsilon_{hi}^\alpha \) and \( |\tilde{A}_{i}^\alpha| \leq \varepsilon_{Ai}^\alpha \) where \( \varepsilon_{hi}^\alpha \) and \( \varepsilon_{Ai}^\alpha \) are sufficiently small, user-defined tolerances.

3. If the model does not produce sustained vibration, human intervention is needed.
Chapter 5
Results

This chapter presents the results for the parameter tuning procedure described in Chapter 4. We applied the tuning process to 14 sets of experimental data extracted from the HSDI system. Four sets are from subjects with healthy voice (labeled H1–H4), two sets are from subjects with polyp (labeled P1 and P2), four sets are from subjects with nodules (labeled N1–N4), and four sets are from subjects unilateral paralysis (labeled U1–U4). In the polyp and unilateral paralysis cases, we labeled the vocal fold with the pathology as the right vocal fold to simplify the data comparison. Before discussing the results, we first describe the system used to collect the experimental data and two data processing techniques to facilitate the parameter tuning process.

5.1 Experimental System

The equipment used to acquire images of the vocal fold vibrations is the KayPEN-TAX High-Speed Video System Model 9700 with 2000 fps and resolution of 120 × 256 pixels, which was described in Section 1.5. The contour of the vocal folds is extracted from each frame of the HSDI video by an image processing algorithm. For a given video, the image was processed as follows. First, the contour of the vocal folds are extracted from each frame of the HSDI video by an image processing algorithm. The glottal axis (i.e., the line connecting the anterior and posterior points) is then determined using first-order linear regression. As shown in Figure 5.1, the displacement of contour point \( i \) on the left vocal fold is defined as perpendicular distance between the point and the glottal axis. By collecting the displacement of all contour points of interest from each frame, the time evolution of the posi-
tion of each contour point is obtained. A picture of one cycle of a normal vocal fold phonation using the just-described procedure is given in Figure 5.2. In this figure, the distances are dimensionless since they were normalized to the largest glottis length (frame 5). It is important to note that during extraction of the vocal fold contour, the image processing algorithm cannot separate the lower and upper edges of the vocal folds [7]. That is, the contour is based on whatever edge (lower or upper) is closest to the glottal axis.

5.2 Data Processing

As described in section 4.1, the goal of the tuning procedure is to match the simulation results from the proposed model to the experimental ones in both the frequency and time domains. This comparison is hampered by two issues: a) the resolution of the standard FFT algorithm is poor, and b) the time-domain curves
from the model and experiment are not necessarily aligned in time. Next, we describe correction procedures for these two issues.

5.2.1 Corrections for FFT Calculation

In spectral analysis, sampled data are transformed into the frequency domain by the discrete Fourier transform (DFT). Direct DFT calculations need $O(N^2)$ arithmetical operations for $N$ points data, and thus is slow for large data sets. In practice, FFT is used to calculate DFT since it significantly reduce the computational effort to $O(N \log N)$ [28]. DFT results have two significant drawbacks [26, 30]. First, it only generates spectral values at a finite number of frequencies $f_i = if_s/N$, where $i = 0, 1, ..., N/2$ for an even $N$ and $f_s$ is sampling frequency. However, peaks might happen between these points, leading to the so-called picket
fence effect or resolution bias error. Note that the frequency resolution of the KayPENTAX High-Speed Video System Model 9700 is about 3.9 Hz since it operates at 2000 fps with \( N = 512 \). Second, since we are not able to exactly sample an integer multiple of periods of the signal, spectral leakage can occur, producing spectral values other than at the signal’s frequency components. To deal with these two problems, several interpolated FFT (IFFT) methods were introduced in [4, 26, 30].

To correct the frequency, amplitude, and phase of the FFT results, we adopted the IFFT method proposed in [26] that employs the barycenter scheme. Referring to Figure 5.3, suppose we found a peak at \( f_k \) with spectrum amplitude \( A_k \) where \( A_{k-1} \) and \( A_{k+1} \) are the spectrum amplitudes at adjacent spectral bins. The corrected peak frequency is given by

\[
f = f_k + \delta f_s/N
\]

where

\[
\delta = \begin{cases} 
-\frac{A_{k-1}}{A_{k-1} + A_k} & \text{if } A_{k-1} > A_{k+1} \\
\frac{A_{k+1}}{A_{k+1} + A_k} & \text{if } A_{k-1} \leq A_{k+1}.
\end{cases}
\]

The corrected peak amplitude is given by

\[
A = \frac{A_k \pi \delta}{\sin \pi \delta} = \frac{A_k}{\text{sinc } \delta},
\]

while the corrected peak phase angle is

\[
\phi = \phi_k + \delta \pi.
\]

5.2.2 Alignment

The model and experimental data need to be aligned in time so they can be properly compared in the time domain. To this end, the time delay between the
model and experimental data has to be estimated. Cross correlation is a widely known technique to calculate the time delay between two related signals. Let $x_e(n)$ be the sampled experimental signal and $x_m(n)$ be the corresponding model signal. Their unbiased cross correlation $R$ is defined as [19]

$$ R(\tau) = \frac{1}{N - \tau} \sum_{n=0}^{N-\tau-1} x_e(n)x_m(n + \tau), \quad \tau = 0, 1, ..., M - 1, \quad M \ll N. \quad (5.5) $$

The time when $R(\tau)$ reaches its first peak is the time delay between $x_e(n)$ and $x_m(n)$. The resolution of the time delay is 0.5 ms as the experimental data is sampled at 2000 Hz. If the vocal folds’ fundamental frequency is 200 Hz and each cycle has 10 sample points, then the phase resolution is about 36°, which is too large for a good alignment. This can be improved by simple parabolic interpolation as shown in Figure 5.4. If the peak $R_{peak}$ and its two neighbors $R_l$ and $R_r$ are fit by the polynomial $R' = a\tau^2 + b\tau + c$, then the time delay $\tau^*$ is obtained when the curve reaches its maximum, i.e.,

$$ \tau^* = -\frac{b}{2a} = \tau_{peak} - \frac{1}{2} \frac{R_r - R_l}{R_r - 2R_{peak} + R_l}. \quad (5.6) $$

39
Figure 5.5 (a) and (b) show an example of the match between experimental and simulation data for $x_2^1(t)$ before and after the alignment procedure was applied.

### 5.3 Analysis of Results

The lists of parameters that resulted from applying the full tuning process of Chapter 4 to the 14 sets of experimental data are shown in Appendix A. The figures comparing the time- and frequency-domain responses of the model simulation to the experimental data for the 14 sets are collected in Appendix B. Before discussing these results, we demonstrate the difference in the model response when only the manual coarse-tuning procedure was applied versus when the automatic fine-tuning procedure was applied. For the purpose of this demonstration, we use the experimental data set for subject H1.
FIGURE 5.5. (a) Experimental and simulation data before alignment. (b) Experimental and simulation data after alignment.

5.3.1 Comparison of Coarse and Fine Tuning

After applying the coarse-tuning procedure to subject data H1, we obtained the following model parameters

\[
\begin{align*}
    m_{\alpha}^\beta_{1,2} &= 0.95, & m_{\alpha}^\beta_{3} &= 0.96, & m_B &= 1.0 \\
    k_{\alpha}^\beta_{1,2} &= 0.64, & k_{\alpha}^\beta_{3} &= 0.66, & k_B &= 1.0, \\
    \kappa_{s1}^\alpha &= 1.7, & \kappa_{s2,3}^\alpha &= 1.1, \\
    \kappa_{\alpha}^\beta_{1,2,3} &= 0.64, & \kappa_{1}^\beta_{4} &= 0.69, & \kappa_{1}^\beta_{4} &= 0.68, \\
    c_{\beta i} &= 0.5k_{\beta i}^\alpha, \\
    d_{l1}^\alpha &= 0.005 \text{ cm}, & d_{l2}^\alpha &= 0.013 \text{ cm}, & d_{l3}^\alpha &= 0.044 \text{ cm}, \\
    d_{u1}^\alpha &= 0.002 \text{ cm}, & d_{u2}^\alpha &= 0.0129 \text{ cm}, & d_{u3}^\alpha &= 0.0439 \text{ cm} \\
    P_s &= 0.012 \text{ g/(cm-ms^2)}. 
\end{align*}
\]

(5.7)

Note that the parameter values without units in (5.7) represent multiplicative factors applied to the corresponding initial values. For example, \( m_{\beta 3}^\alpha = 0.96 \) is 96% of...
the parameter value given in (3.10). 280 ms simulation based on the coarse tuning parameter set (3.10) is shown in Figures B.1 and B.2. Figure B.1 compares the time domain behavior of the variables \( x_i^\alpha \) from the model simulation and experiment while Figure B.2 compares the FFT of \( x_i^\alpha \). Only a 35 ms portion of the time response is displayed in B.1 to facilitate the visualization. For comparison purposes, the error variables defined in Section 4.2 were calculated for the parameter set in (5.7) yielding the following values

\[
\left| \tilde{f} \right| = 3.06 \, \text{Hz}, \quad \max_{\alpha, i} \left\{ \left| \tilde{h}_i^\alpha \right| \right\} = 0.0171 \, \text{cm}, \quad \max_{\alpha, i} \left\{ \left| \tilde{A}_i^\alpha \right| \right\} = 0.0147 \, \text{cm}. \quad (5.8)
\]

The automatic fine-tuning procedure was then applied, starting with the parameters in (5.7). The tolerances were set to \( \varepsilon_f = 1 \, \text{Hz} \), \( \varepsilon_{h_i}^\alpha = 0.01 \, \text{cm} \), and \( \varepsilon_{A_i}^\alpha = 0.01 \, \text{cm} \). As a result, the following updated parameters were obtained

\[
\begin{align*}
m_{l, \beta 1, 2}^1 &= 0.936, & m_{l, \beta 3}^1 &= 0.989, \\
m_{r, \beta 1, 2}^1 &= 0.936, & m_{r, \beta 3}^1 &= 0.945, \\
k_{l, \beta 1, 2}^1 &= 0.650, & k_{l, \beta 3}^1 &= 0.701, \\
k_{r, \beta 1}^1 &= 0.650, & k_{r, \beta 3}^1 &= 0.670, \\
\kappa_{l, \beta 1, 2, 3}^1 &= 0.650, & \kappa_{l, \beta 4}^1 &= 0.733, \\
\kappa_{r, \beta 1, 2, 3}^1 &= 0.650, & \kappa_{r, \beta 4}^1 &= 0.690.
\end{align*}
\]

(5.9)

The fine-tuning results are shown in Figures B.3 and B.4. After the automatic fine tuning, the errors became

\[
\left| \tilde{f} \right| = 0.220 \, \text{Hz}, \quad \max_{\alpha, i} \left\{ \left| \tilde{h}_i^\alpha \right| \right\} = 0.0099 \, \text{cm}, \quad \max_{\alpha, i} \left\{ \left| \tilde{A}_i^\alpha \right| \right\} = 0.0089 \, \text{cm}. \quad (5.10)
\]

In comparison to (5.8), this represents a reduction of approximately 93% in \( \tilde{f} \), 42% in \( \max_{\alpha, i} \left\{ \left| \tilde{h}_i^\alpha \right| \right\} \), and 40% in \( \max_{\alpha, i} \left\{ \left| \tilde{A}_i^\alpha \right| \right\} \).

### 5.3.2 Discussion

Recall that step 3 of the course-tuning procedure is devoted to matching the duration of collisions between the left and right vocal folds. Due to coupled nature of
the model, tuning the collision duration in one segment of the vocal fold affected that of other segments. For example, increasing the collision duration in the ventral segment by tuning $d_{\beta_1}^\alpha$, $c_{\kappa_1}$, and $c_{\alpha_1}$ would also increase the collision duration in the dorsal segment. Therefore, it was very difficult to match the collision duration in all segments so a compromise had to be reached. For example, in Figure B.1, the collision duration in the ventral segment in the simulation is smaller than in the experiment, while in the dorsal segment it is larger.

From the experimental curves in Figures B.1, B.5, B.7, and B.9, we can see that the healthy subjects exhibit relatively good symmetry and little phase difference between the left and right vocal folds. Therefore, it was relatively easy to tune the model parameters for these cases. The resulting parameters had smaller variation from right to left and from the anterior segment to the posterior segment compared to the pathological cases (this is quantified below). Further, the biomechanical model was able to maintain sustained vibrations in a large range of the parameter space, allowing the fine tuning algorithm to finish without human intervention. In the pathological cases, the model produced sustained vibrations only in a narrower region of parameter space, causing the automatic fine-tuning process to terminate. In these situations, we manually tuned the parameters following the procedure of the fine-tuning algorithm.

In the polyp cases, subject P1 has a small polyp near the ventral segment of the right vocal fold, while subject P2 has a large polyp near the medial segment of the right vocal fold. In both cases, the polyp segment of the right vocal fold exhibited a smaller displacement amplitude and lagged the corresponding healthy segment of the left vocal fold; see experimental curves in Figures B.11 and B.13. The smaller displacement is due to the fact that the collision of the right and left vocal folds in the polyp segment occurred to the left of the glottal axis in both P1 and P2.
The model and tuning process captured the above phenomena by producing larger values for the mass and/or stiffnesses associated with the polyp segment; see values of \( m_{\beta_1} \), \( k_{\beta_1} \), and \( \kappa_{\beta_1} \) for subject P1, and value for \( m_{\beta_2} \) for subject P2 in Table A.2.

Subjects N1~N4 have a nodule on both vocal folds. Subjects N1 and N4 have a swelling at the ventral segment of the right and left vocal folds, while subjects N2 and N3 exhibit a swelling at the medial segment. As a result, the model produced larger mass and/or stiffness at the nodule segments; see values of \( m_{\alpha_1} \), \( k_{\alpha_1} \), and \( \kappa_{\alpha_1} \) for subjects N1 and N4, and values of \( m_{\alpha_2} \), \( k_{\alpha_2} \), and \( \kappa_{\alpha_2} \) for subjects N2 and N3 in Table A.3.

The unilateral paralysis cases have a large phase difference between the displacement of the left and right vocal folds with the healthy side leading the paralyzed side. The tuning results produced larger cover masses for the paralyzed (right) side, especially for subjects U1~U3; see values of \( m_{\beta_1} \) in Table A.4. This phenomenon however is not apparent from the images of the HSDI system. A physiological explanation for the paralyzed vocal fold having a larger cover mass is that, without the muscle motor activity, the arytenoid cartilage cannot properly hold the vocal folds in place and thus a deeper layer of the vocal fold vibrates which makes the effective cover mass larger.

Asymmetries between the left and right vocal folds and anterior and posterior segments of the vocal folds are important indicators of the existence of a vocal fold pathology. Next, we will introduce model parameter-dependent measures for these asymmetries in order to relate regions of the parameter space to the studied pathologies. The proposed measures of asymmetry are based on the cover mass \( m_{\alpha_1} \) and stiffness between body and cover masses \( k_{\beta_1} \). Specifically, we define

\[
Q_{l/r} = \max \left\{ Q_{l/r}^m, Q_{l/r}^k \right\},
\]

(5.11)
where

\[
Q_{l/r}^m = \frac{\max_{\alpha} \left\{ \sum_{i,\beta} m_{\beta i}^\alpha \right\}}{\min_{\alpha} \left\{ \sum_{i,\beta} m_{\beta i}^\alpha \right\}} \quad \text{and} \quad Q_{l/r}^k = \frac{\max_{\alpha} \left\{ \sum_{i,\beta} k_{\beta i}^\alpha \right\}}{\min_{\alpha} \left\{ \sum_{i,\beta} k_{\beta i}^\alpha \right\}},
\]

(5.12)
to estimate the left/right asymmetry. We define

\[
Q_{a/p} = \max \left\{ Q_{a/p}^m, Q_{a/p}^k \right\},
\]

(5.13)
where

\[
Q_{a/p}^m = \max \left\{ \frac{\max_{i,\beta} \left\{ m_{\beta i}^\alpha \right\}}{\min_{i,\beta} \left\{ m_{\beta i}^\alpha \right\}}, \frac{\max_{i,\beta} \left\{ m_{\beta i}^\alpha \right\}}{\min_{i,\beta} \left\{ m_{\beta i}^\alpha \right\}} \right\}
\]

and

\[
Q_{a/p}^k = \max \left\{ \frac{\max_{i,\beta} \left\{ k_{\beta i}^\alpha \right\}}{\min_{i,\beta} \left\{ k_{\beta i}^\alpha \right\}}, \frac{\max_{i,\beta} \left\{ k_{\beta i}^\alpha \right\}}{\min_{i,\beta} \left\{ k_{\beta i}^\alpha \right\}} \right\},
\]

(5.14)
to estimate the anterior/posterior asymmetry. Note from the definitions in (5.12) and (5.14) that \(Q_{l/r} \geq 1\) and \(Q_{a/p} \geq 1\). For healthy vocal folds, we expect both factors to be approximately one [8, 27].

Based on the model parameters in Appendix A, we calculated the above factors for each of the 14 subjects. The results are shown in Table 5.1. We then generated a plot of \(Q_{l/r}\) versus \(Q_{a/p}\) to aid in correlating the asymmetry plane to each pathology; see Figure 5.6. As expected, the asymmetry factors for the healthy subjects are the closest to bottom left corner of the asymmetry plane, i.e., where \(Q_{l/r} \approx 1\) and \(Q_{a/p} \approx 1\). Nodule cases have good left/right symmetry but large anterior/posterior asymmetry, and thus are located immediately above the healthy cases in the asymmetry plane. There is a small intersection between the healthy and nodule regions, which is not unexpected since small nodules have little influence on vocal fold vibration [5]. Unilateral paralysis mainly have large left/right asymmetry but normal anterior/posterior asymmetry. Thus, its region in the asymmetry plane is located immediately to the right of the healthy cases. Finally, polyp cases showed a large
TABLE 5.1. Asymmetry parameters.

<table>
<thead>
<tr>
<th></th>
<th>$Q_{l/r}^m$</th>
<th>$Q_{l/r}^k$</th>
<th>$Q_{l/r}^l$</th>
<th>$Q_{a/p}^k$</th>
<th>$Q_{a/p}^k$</th>
<th>$Q_{a/p}^l$</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>1.015</td>
<td>1.016</td>
<td>1.016</td>
<td>1.057</td>
<td>1.079</td>
<td>1.079</td>
</tr>
<tr>
<td>H2</td>
<td>1.024</td>
<td>1.065</td>
<td>1.065</td>
<td>1.126</td>
<td>1.135</td>
<td>1.135</td>
</tr>
<tr>
<td>H3</td>
<td>1.111</td>
<td>1.018</td>
<td>1.111</td>
<td>1.157</td>
<td>1.207</td>
<td>1.207</td>
</tr>
<tr>
<td>H4</td>
<td>1.055</td>
<td>1.034</td>
<td>1.055</td>
<td>1.190</td>
<td>1.144</td>
<td>1.190</td>
</tr>
<tr>
<td>P1</td>
<td>1.167</td>
<td>1.081</td>
<td>1.167</td>
<td>1.500</td>
<td>1.200</td>
<td>1.500</td>
</tr>
<tr>
<td>P2</td>
<td>1.233</td>
<td>1.143</td>
<td>1.233</td>
<td>1.700</td>
<td>1.429</td>
<td>1.700</td>
</tr>
<tr>
<td>N1</td>
<td>1.058</td>
<td>1.133</td>
<td>1.133</td>
<td>1.233</td>
<td>1.222</td>
<td>1.233</td>
</tr>
<tr>
<td>N2</td>
<td>1.064</td>
<td>1.137</td>
<td>1.137</td>
<td>1.442</td>
<td>1.849</td>
<td>1.849</td>
</tr>
<tr>
<td>N3</td>
<td>1.070</td>
<td>1.018</td>
<td>1.070</td>
<td>1.674</td>
<td>1.846</td>
<td>1.846</td>
</tr>
<tr>
<td>N4</td>
<td>1.041</td>
<td>1.053</td>
<td>1.053</td>
<td>1.575</td>
<td>1.364</td>
<td>1.575</td>
</tr>
<tr>
<td>U1</td>
<td>1.140</td>
<td>1.000</td>
<td>1.140</td>
<td>1.022</td>
<td>1.045</td>
<td>1.045</td>
</tr>
<tr>
<td>U2</td>
<td>1.101</td>
<td>1.199</td>
<td>1.199</td>
<td>1.000</td>
<td>1.020</td>
<td>1.020</td>
</tr>
<tr>
<td>U3</td>
<td>1.241</td>
<td>1.000</td>
<td>1.241</td>
<td>1.015</td>
<td>1.050</td>
<td>1.050</td>
</tr>
<tr>
<td>U4</td>
<td>1.002</td>
<td>1.667</td>
<td>1.667</td>
<td>1.013</td>
<td>1.000</td>
<td>1.013</td>
</tr>
</tbody>
</table>

anterior/posterior asymmetry, compared to the nodule, and moderate left/right asymmetry. Their region lies above the $Q_{l/r} = Q_{a/p}$ line and to the right of the nodule region.
FIGURE 5.6. Plot of \( Q_{l/r} \) versus \( Q_{a/p} \) and classification of asymmetry plane.
Chapter 6
Conclusion and Recommendations

Vibration of vocal folds is the main source of voice production. Irregularity in the sustained vibration of vocal folds is believed to be rooted in physiological changes in the vocal folds caused by certain disorders. This dissertation introduced a new biomechanical model for the vibrations of vocal folds that incorporates the most relevant characteristics of the vocal fold dynamics. The proposed lumped-parameter model utilizes one mass for the vocal fold body tissue and three pairs of upper/lower masses along the vocal fold length for the cover tissue. A preliminary model validation was presented by evaluating its ability to reproduce the actual vibration pattern of healthy and pathological (nodule, polyp, and unilateral paralysis) vocal folds obtained experimentally via a high-speed imaging system. The proposed model was able to match the experimental (frequency- and time-domain) data relatively well while outperforming an existing model that neglects the mass and stiffness of the vocal fold body tissue.

Another contribution of this work was the development of a physically-inspired, semi-empirical tuning procedure for the model parameters. The tuning criteria were based on the fundamental frequency, amplitude, phase difference, and collision duration of the vocal fold vibration. Finally, we introduced two factors to quantify the left/right and anterior/posterior asymmetries of vocal folds. The use of only two factors facilitates classifying and visualizing how the healthy and pathological vocal folds relate to these asymmetries. That is, depending on where these factors are located on the asymmetry plane for a given set of model parameters, one can tell if the vocal folds are healthy or not and what disorder (nodule, polyp, or
unilateral paralysis) is present. This ‘chart’ can potentially further the assessment and diagnosis of voice disorders in clinical settings.

Our recommendations for future work are as follows:

• The results of this dissertation are based on experimental data from a limited set of subjects. Therefore, a more detailed study should be conducted using a larger and statistically-meaningful number of subjects.

• The experimental data was captured by a HSDI system with 2000 fps and resolution of $120 \times 250$ pixels. Since the accuracy of the tuning results for the model parameters is affected by these characteristics, the use of a HSDI system with higher frame rate and resolution is recommended for future research.

• The proposed biomechanical model incorporates a commonly-used model for the collision of the vocal folds. We observed that the model cannot be adequately tuned to capture the duration of the collision at each segment of the vocal fold. Therefore, further research efforts should be dedicated to improving the collision model.

• The fine-tuning algorithm may be improved by introducing and minimizing other error variables.
References


Appendix A
Parameter Tuning Results

The following tables compile the parameters that resulted from applying the full tuning process of Chapter 4 to the 14 sets of experimental data. The parameter values without units in the tables represent multiplicative factors applied to the corresponding initial values. For example, \( m_{\beta_1}^1 = 0.936 \) in the first column of Table A.1 means 93.6% of the corresponding parameter value given in (3.10).
<table>
<thead>
<tr>
<th></th>
<th>H1</th>
<th>H2</th>
<th>H3</th>
<th>H4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m^i_{\beta 1}$</td>
<td>0.936</td>
<td>0.780</td>
<td>0.707</td>
<td>0.637</td>
</tr>
<tr>
<td>$m^i_{\beta 2}$</td>
<td>0.936</td>
<td>0.791</td>
<td>0.707</td>
<td>0.569</td>
</tr>
<tr>
<td>$m^i_{\beta 3}$</td>
<td>0.989</td>
<td>0.878</td>
<td>0.818</td>
<td>0.676</td>
</tr>
<tr>
<td>$m^r_{\beta 1}$</td>
<td>0.936</td>
<td>0.760</td>
<td>0.646</td>
<td>0.598</td>
</tr>
<tr>
<td>$m^r_{\beta 2}$</td>
<td>0.936</td>
<td>0.780</td>
<td>0.646</td>
<td>0.549</td>
</tr>
<tr>
<td>$m^r_{\beta 3}$</td>
<td>0.946</td>
<td>0.853</td>
<td>0.717</td>
<td>0.637</td>
</tr>
<tr>
<td>$k^i_{\beta 1}$</td>
<td>0.650</td>
<td>1.192</td>
<td>0.861</td>
<td>1.255</td>
</tr>
<tr>
<td>$k^i_{\beta 2}$</td>
<td>0.650</td>
<td>1.192</td>
<td>0.960</td>
<td>1.347</td>
</tr>
<tr>
<td>$k^i_{\beta 3}$</td>
<td>0.701</td>
<td>1.352</td>
<td>0.980</td>
<td>1.347</td>
</tr>
<tr>
<td>$k^r_{\beta 1}$</td>
<td>0.650</td>
<td>1.153</td>
<td>0.812</td>
<td>1.275</td>
</tr>
<tr>
<td>$k^r_{\beta 2}$</td>
<td>0.650</td>
<td>1.153</td>
<td>0.960</td>
<td>1.347</td>
</tr>
<tr>
<td>$k^r_{\beta 3}$</td>
<td>0.670</td>
<td>1.201</td>
<td>0.980</td>
<td>1.459</td>
</tr>
<tr>
<td>$\kappa^\alpha_{s 1}$</td>
<td>1.7</td>
<td>1.1</td>
<td>1.3</td>
<td>2.1</td>
</tr>
<tr>
<td>$\kappa^\alpha_{s 2,3}$</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>1.6</td>
</tr>
<tr>
<td>$\kappa^i_{\beta 1}$</td>
<td>0.650</td>
<td>1.192</td>
<td>0.861</td>
<td>1.030</td>
</tr>
<tr>
<td>$\kappa^i_{\beta 2,3}$</td>
<td>0.650</td>
<td>1.192</td>
<td>1.257</td>
<td>1.357</td>
</tr>
<tr>
<td>$\kappa^i_{\beta 4}$</td>
<td>0.733</td>
<td>2.414</td>
<td>2.47</td>
<td>2.142</td>
</tr>
<tr>
<td>$\kappa^r_{\beta 1}$</td>
<td>0.650</td>
<td>1.153</td>
<td>0.822</td>
<td>1.030</td>
</tr>
<tr>
<td>$\kappa^r_{\beta 2,3}$</td>
<td>0.650</td>
<td>1.153</td>
<td>0.960</td>
<td>1.357</td>
</tr>
<tr>
<td>$\kappa^r_{\beta 4}$</td>
<td>0.690</td>
<td>2.210</td>
<td>2.347</td>
<td>2.346</td>
</tr>
<tr>
<td>$c_{\beta i}$</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$d^i_{l 1}$ (cm)</td>
<td>0.005</td>
<td>0.015</td>
<td>0.03</td>
<td>0.025</td>
</tr>
<tr>
<td>$d^i_{l 2}$ (cm)</td>
<td>0.013</td>
<td>0.028</td>
<td>0.040</td>
<td>0.028</td>
</tr>
<tr>
<td>$d^i_{l 3}$ (cm)</td>
<td>0.044</td>
<td>0.049</td>
<td>0.054</td>
<td>0.044</td>
</tr>
<tr>
<td>$d^i_{r 1}$ (cm)</td>
<td>0.005</td>
<td>0.020</td>
<td>0.010</td>
<td>0.015</td>
</tr>
<tr>
<td>$d^i_{r 2}$ (cm)</td>
<td>0.013</td>
<td>0.038</td>
<td>0.036</td>
<td>0.008</td>
</tr>
<tr>
<td>$d^i_{r 3}$ (cm)</td>
<td>0.044</td>
<td>0.074</td>
<td>0.064</td>
<td>0.044</td>
</tr>
<tr>
<td>$d^r_{n 1}$ (cm)</td>
<td>0.0019</td>
<td>0.0119</td>
<td>0.0269</td>
<td>0.0219</td>
</tr>
<tr>
<td>$d^r_{n 2}$ (cm)</td>
<td>0.0129</td>
<td>0.0279</td>
<td>0.0399</td>
<td>0.0229</td>
</tr>
<tr>
<td>$d^r_{n 3}$ (cm)</td>
<td>0.0439</td>
<td>0.0489</td>
<td>0.0539</td>
<td>0.0439</td>
</tr>
<tr>
<td>$d^r_{n 4}$ (cm)</td>
<td>0.0019</td>
<td>0.0169</td>
<td>0.0069</td>
<td>0.0119</td>
</tr>
<tr>
<td>$d^r_{n 5}$ (cm)</td>
<td>0.0129</td>
<td>0.0379</td>
<td>0.0539</td>
<td>0.0229</td>
</tr>
<tr>
<td>$d^r_{n 6}$ (cm)</td>
<td>0.0439</td>
<td>0.0739</td>
<td>0.0639</td>
<td>0.0439</td>
</tr>
<tr>
<td>$\zeta_1$ (no collision)</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>$\zeta_1$ (collision)</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>$P_s$ ($\frac{g}{cm \cdot ms^2}$)</td>
<td>0.012</td>
<td>0.012</td>
<td>0.010</td>
<td>0.013</td>
</tr>
<tr>
<td>Parameter</td>
<td>P1</td>
<td>P2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>------</td>
<td>------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$m_{p1}$</td>
<td>0.560</td>
<td>1.188</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$m_{p2}$</td>
<td>0.560</td>
<td>1.188</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$m_{p3}$</td>
<td>0.560</td>
<td>1.188</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$m_{p1}$</td>
<td>0.840</td>
<td>1.188</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$m_{p2}$</td>
<td>0.560</td>
<td>2.020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$m_{p3}$</td>
<td>0.560</td>
<td>1.888</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{a1}^1$</td>
<td>0.864</td>
<td>1.170</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{a2}^1$</td>
<td>0.864</td>
<td>1.170</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{a3}^1$</td>
<td>0.864</td>
<td>0.819</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{a1}^2$</td>
<td>1.051</td>
<td>1.170</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{a2}^2$</td>
<td>0.876</td>
<td>1.270</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{a3}^2$</td>
<td>0.876</td>
<td>1.170</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{a1}^3$</td>
<td>1.1</td>
<td>1.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{a2}^3$</td>
<td>0.864</td>
<td>1.053</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{a3}^3$</td>
<td>0.864</td>
<td>1.170</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$c_{a3}$</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$d_{l1}$ (cm)</td>
<td>0.015</td>
<td>0.047</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$d_{l2}$ (cm)</td>
<td>0.018</td>
<td>0.028</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$d_{l3}$ (cm)</td>
<td>0.054</td>
<td>0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$d_{u1}$ (cm)</td>
<td>-0.005</td>
<td>0.057</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$d_{u2}$ (cm)</td>
<td>0.018</td>
<td>-0.032</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$d_{u3}$ (cm)</td>
<td>0.074</td>
<td>0.054</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$d_{l1}$ (cm)</td>
<td>0.0119</td>
<td>0.0419</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$d_{l2}$ (cm)</td>
<td>0.0179</td>
<td>0.0279</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$d_{l3}$ (cm)</td>
<td>0.0639</td>
<td>0.0139</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$d_{u1}$ (cm)</td>
<td>-0.0081</td>
<td>0.0519</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$d_{u2}$ (cm)</td>
<td>0.0179</td>
<td>-0.032</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$d_{u3}$ (cm)</td>
<td>0.0839</td>
<td>0.0439</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\zeta_1$ (no collision)</td>
<td>0.35</td>
<td>0.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\zeta_1$ (collision)</td>
<td>0.7</td>
<td>0.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$P_s$ ($\frac{g}{\text{cm}-\text{ms}^2}$)</td>
<td>0.013</td>
<td>0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parameters</td>
<td>N1</td>
<td>N2</td>
<td>N3</td>
<td>N4</td>
</tr>
<tr>
<td>---------------------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>$m_{\beta_1}$</td>
<td>0.53</td>
<td>0.54</td>
<td>0.44</td>
<td>1.37</td>
</tr>
<tr>
<td>$m_{\beta_2}$</td>
<td>0.43</td>
<td>0.65</td>
<td>0.68</td>
<td>0.97</td>
</tr>
<tr>
<td>$m_{\beta_3}$</td>
<td>0.43</td>
<td>0.54</td>
<td>0.46</td>
<td>0.87</td>
</tr>
<tr>
<td>$m_{\beta_1}$</td>
<td>0.55</td>
<td>0.57</td>
<td>0.46</td>
<td>1.37</td>
</tr>
<tr>
<td>$m_{\beta_2}$</td>
<td>0.46</td>
<td>0.75</td>
<td>0.77</td>
<td>0.97</td>
</tr>
<tr>
<td>$m_{\beta_3}$</td>
<td>0.46</td>
<td>0.52</td>
<td>0.46</td>
<td>1.00</td>
</tr>
<tr>
<td>$k_{\beta_1}$</td>
<td>1.10</td>
<td>0.83</td>
<td>1.30</td>
<td>1.50</td>
</tr>
<tr>
<td>$k_{\beta_2}$</td>
<td>1.20</td>
<td>1.35</td>
<td>2.40</td>
<td>1.20</td>
</tr>
<tr>
<td>$k_{\beta_3}$</td>
<td>1.10</td>
<td>0.73</td>
<td>1.90</td>
<td>1.10</td>
</tr>
<tr>
<td>$k_{\beta_1}$</td>
<td>1.10</td>
<td>0.93</td>
<td>1.30</td>
<td>1.50</td>
</tr>
<tr>
<td>$k_{\beta_2}$</td>
<td>1.00</td>
<td>1.35</td>
<td>2.30</td>
<td>1.20</td>
</tr>
<tr>
<td>$k_{\beta_3}$</td>
<td>0.90</td>
<td>1.03</td>
<td>1.90</td>
<td>1.30</td>
</tr>
<tr>
<td>$\kappa_{si}$</td>
<td>1.1</td>
<td>1.0</td>
<td>1.6</td>
<td>1.00</td>
</tr>
<tr>
<td>$\kappa_{\beta_1}$</td>
<td>1.10</td>
<td>0.94</td>
<td>1.00</td>
<td>1.50</td>
</tr>
<tr>
<td>$\kappa_{\beta_2,3}$</td>
<td>1.10</td>
<td>1.30</td>
<td>2.50</td>
<td>1.20</td>
</tr>
<tr>
<td>$\kappa_{\beta_4}$</td>
<td>1.10</td>
<td>0.70</td>
<td>1.60</td>
<td>1.00</td>
</tr>
<tr>
<td>$\kappa_{\beta_1}$</td>
<td>1.10</td>
<td>1.05</td>
<td>0.90</td>
<td>1.50</td>
</tr>
<tr>
<td>$\kappa_{\beta_2,3}$</td>
<td>1.0</td>
<td>1.30</td>
<td>2.40</td>
<td>1.20</td>
</tr>
<tr>
<td>$\kappa_{\beta_4}$</td>
<td>0.90</td>
<td>1.04</td>
<td>1.60</td>
<td>1.30</td>
</tr>
<tr>
<td>$c_{\beta_1}$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$d_{l_{1}}$ (cm)</td>
<td>0.005</td>
<td>0.025</td>
<td>0.017</td>
<td>0.005</td>
</tr>
<tr>
<td>$d_{l_{2}}$ (cm)</td>
<td>0.018</td>
<td>0.033</td>
<td>0.013</td>
<td>-0.002</td>
</tr>
<tr>
<td>$d_{l_{3}}$ (cm)</td>
<td>0.051</td>
<td>0.044</td>
<td>0.022</td>
<td>0.074</td>
</tr>
<tr>
<td>$d_{l_{1}}$ (cm)</td>
<td>0.015</td>
<td>0.045</td>
<td>0.017</td>
<td>0.005</td>
</tr>
<tr>
<td>$d_{l_{2}}$ (cm)</td>
<td>0.018</td>
<td>-0.020</td>
<td>0.018</td>
<td>0.038</td>
</tr>
<tr>
<td>$d_{l_{3}}$ (cm)</td>
<td>0.031</td>
<td>0.074</td>
<td>0.022</td>
<td>0.094</td>
</tr>
<tr>
<td>$d_{u_{1}}$ (cm)</td>
<td>0.0019</td>
<td>0.0219</td>
<td>0.0139</td>
<td>0.0019</td>
</tr>
<tr>
<td>$d_{u_{2}}$ (cm)</td>
<td>0.0179</td>
<td>0.033</td>
<td>0.0179</td>
<td>-0.0021</td>
</tr>
<tr>
<td>$d_{u_{3}}$ (cm)</td>
<td>0.0509</td>
<td>0.0439</td>
<td>0.0219</td>
<td>0.0639</td>
</tr>
<tr>
<td>$d_{u_{1}}$ (cm)</td>
<td>0.0069</td>
<td>0.419</td>
<td>0.0139</td>
<td>0.0119</td>
</tr>
<tr>
<td>$d_{u_{2}}$ (cm)</td>
<td>0.0179</td>
<td>-0.0200</td>
<td>0.0179</td>
<td>0.0379</td>
</tr>
<tr>
<td>$d_{u_{3}}$ (cm)</td>
<td>0.03087</td>
<td>0.0739</td>
<td>0.0219</td>
<td>0.1039</td>
</tr>
<tr>
<td>$\zeta_{1}$ (no collision)</td>
<td>0.30</td>
<td>0.35</td>
<td>0.40</td>
<td>0.40</td>
</tr>
<tr>
<td>$\zeta_{1}$ (collision)</td>
<td>0.6</td>
<td>0.7</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>$P_s$ ($\frac{g}{cm\cdot ms^2}$)</td>
<td>0.011</td>
<td>0.013</td>
<td>0.010</td>
<td>0.013</td>
</tr>
<tr>
<td>Parameter</td>
<td>U1</td>
<td>U2</td>
<td>U3</td>
<td>U4</td>
</tr>
<tr>
<td>-----------</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>$m_{\beta 1}^1$</td>
<td>1.488</td>
<td>2.18</td>
<td>0.70</td>
<td>1.52</td>
</tr>
<tr>
<td>$m_{\beta 2}^1$</td>
<td>1.520</td>
<td>2.18</td>
<td>0.70</td>
<td>1.52</td>
</tr>
<tr>
<td>$m_{\beta 3}^1$</td>
<td>1.488</td>
<td>2.18</td>
<td>0.70</td>
<td>1.53</td>
</tr>
<tr>
<td>$m_{\beta 1}^2$</td>
<td>1.720</td>
<td>2.40</td>
<td>0.87</td>
<td>1.52</td>
</tr>
<tr>
<td>$m_{\beta 2}^2$</td>
<td>1.684</td>
<td>2.40</td>
<td>0.87</td>
<td>1.52</td>
</tr>
<tr>
<td>$m_{\beta 3}^2$</td>
<td>1.720</td>
<td>2.40</td>
<td>0.88</td>
<td>1.54</td>
</tr>
<tr>
<td>$k_{\beta 1}^1$</td>
<td>0.770</td>
<td>0.60</td>
<td>1.47</td>
<td>0.80</td>
</tr>
<tr>
<td>$k_{\beta 2}^1$</td>
<td>0.805</td>
<td>0.60</td>
<td>1.47</td>
<td>0.80</td>
</tr>
<tr>
<td>$k_{\beta 3}^1$</td>
<td>0.770</td>
<td>0.61</td>
<td>1.54</td>
<td>0.80</td>
</tr>
<tr>
<td>$k_{\beta 1}^2$</td>
<td>0.770</td>
<td>0.50</td>
<td>1.47</td>
<td>0.48</td>
</tr>
<tr>
<td>$k_{\beta 2}^2$</td>
<td>0.805</td>
<td>0.50</td>
<td>1.47</td>
<td>0.48</td>
</tr>
<tr>
<td>$k_{\beta 3}^2$</td>
<td>0.770</td>
<td>0.51</td>
<td>1.54</td>
<td>0.48</td>
</tr>
<tr>
<td>$\kappa_{\beta 1}^1$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.5</td>
<td>0.90</td>
</tr>
<tr>
<td>$\kappa_{\beta 2}^1$</td>
<td>1.2</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$\kappa_{\beta 3}^1$</td>
<td>0.60</td>
<td>0.60</td>
<td>1.47</td>
<td>0.80</td>
</tr>
<tr>
<td>$\kappa_{\beta 1}^2$</td>
<td>0.63</td>
<td>0.60</td>
<td>1.47</td>
<td>0.80</td>
</tr>
<tr>
<td>$\kappa_{\beta 2}^2$</td>
<td>0.60</td>
<td>0.61</td>
<td>1.54</td>
<td>0.80</td>
</tr>
<tr>
<td>$\kappa_{\beta 3}^2$</td>
<td>0.60</td>
<td>0.55</td>
<td>1.47</td>
<td>0.48</td>
</tr>
<tr>
<td>$\kappa_{\beta 1}^3$</td>
<td>0.63</td>
<td>0.55</td>
<td>1.47</td>
<td>0.48</td>
</tr>
<tr>
<td>$\kappa_{\beta 2}^3$</td>
<td>0.60</td>
<td>0.56</td>
<td>1.54</td>
<td>0.48</td>
</tr>
<tr>
<td>$\kappa_{\beta 3}^3$</td>
<td>0.60</td>
<td>0.56</td>
<td>1.54</td>
<td>0.48</td>
</tr>
<tr>
<td>$c_{\beta i}$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$d_{\beta 1}^1$ (cm)</td>
<td>$-0.005$</td>
<td>$0.015$</td>
<td>$0.005$</td>
<td>$0.090$</td>
</tr>
<tr>
<td>$d_{\beta 2}^1$ (cm)</td>
<td>$-0.002$</td>
<td>$-0.002$</td>
<td>$0.018$</td>
<td>$0.103$</td>
</tr>
<tr>
<td>$d_{\beta 3}^1$ (cm)</td>
<td>$-0.011$</td>
<td>$0.054$</td>
<td>$0.059$</td>
<td>$0.094$</td>
</tr>
<tr>
<td>$d_{\beta 1}^2$ (cm)</td>
<td>$0.035$</td>
<td>$0.015$</td>
<td>$0.005$</td>
<td>$0.015$</td>
</tr>
<tr>
<td>$d_{\beta 2}^2$ (cm)</td>
<td>$0.018$</td>
<td>$0.028$</td>
<td>$0.018$</td>
<td>$0.028$</td>
</tr>
<tr>
<td>$d_{\beta 3}^2$ (cm)</td>
<td>$-0.001$</td>
<td>$0.004$</td>
<td>$0.059$</td>
<td>$0.024$</td>
</tr>
<tr>
<td>$d_{\beta 1}^3$ (cm)</td>
<td>$-0.0081$</td>
<td>$0.0019$</td>
<td>$0.0119$</td>
<td>$0.0869$</td>
</tr>
<tr>
<td>$d_{\beta 2}^3$ (cm)</td>
<td>$-0.0021$</td>
<td>$-0.0021$</td>
<td>$0.0179$</td>
<td>$0.1029$</td>
</tr>
<tr>
<td>$d_{\beta 3}^3$ (cm)</td>
<td>$-0.0111$</td>
<td>$0.0539$</td>
<td>$0.0589$</td>
<td>$0.0939$</td>
</tr>
<tr>
<td>$d_{\beta 1}^4$ (cm)</td>
<td>$0.0319$</td>
<td>$0.0119$</td>
<td>$0.0119$</td>
<td>$0.0119$</td>
</tr>
<tr>
<td>$d_{\beta 2}^4$ (cm)</td>
<td>$0.0179$</td>
<td>$0.0279$</td>
<td>$0.0199$</td>
<td>$0.0279$</td>
</tr>
<tr>
<td>$d_{\beta 3}^4$ (cm)</td>
<td>$-0.0061$</td>
<td>$0.0039$</td>
<td>$0.0589$</td>
<td>$0.0239$</td>
</tr>
<tr>
<td>$\zeta_1$ (no collision)</td>
<td>0.30</td>
<td>0.40</td>
<td>0.30</td>
<td>0.35</td>
</tr>
<tr>
<td>$\zeta_1$ (collision)</td>
<td>0.6</td>
<td>0.8</td>
<td>0.6</td>
<td>0.7</td>
</tr>
<tr>
<td>$P_s$ ($\frac{\text{g}}{\text{cm-ms}^2}$)</td>
<td>0.012</td>
<td>0.010</td>
<td>0.015</td>
<td>0.006</td>
</tr>
</tbody>
</table>

57
Appendix B
Tuning Results Figures

The following figures depict the time and frequency responses of the model and experiment for the 14 subjects. We note that the portions of the time response of \( x_1^\alpha \) where the left and right curves are indistinguishable denote collision of the \( i \)th segment of the right and left vocal folds.

FIGURE B.3. Automatic fine tuning for subject H1: simulation and experimental time responses.

FIGURE B.4. Automatic fine tuning for subject H1: simulation and experimental amplitude spectra.


FIGURE B.10. Fine tuning for subject H4: simulation and experimental amplitude spectra.

FIGURE B.12. Fine tuning for subject P1: simulation and experimental amplitude spectra.


FIGURE B.18. Fine tuning for subject N2: simulation and experimental amplitude spectra.


FIGURE B.22. Fine tuning for subject N4: simulation and experimental amplitude spectra.


FIGURE B.27. Fine tuning for subject U3: simulation and experimental time responses.


FIGURE B.30. Fine tuning for subject U4: simulation and experimental amplitude spectra.
Vita

Zhenyi Wei was born in 1978, in Zhejiang Province, China. He received a Bachelor of Science Degree in Mechanical Engineering from Zhejiang University, China in June 2001. He earned a master of science degree in control science and engineering from the same university in June 2004. In August 2007 he came to Louisiana State University and since then worked with Dr. Marcio de Queiroz as a research assistant. He is currently a candidate for the degree of Doctor of Philosophy in mechanical engineering.