Signal generation by an uncertain nonlinear dynamical system: application to the production of voiced sounds and experimental validation

E. Cataldo, Christian Soize, R. Sampaio, Christophe Desceliers

Signal generation by an uncertain nonlinear dynamical system: application to the production of voiced sounds and experimental validation.

Edson Cataldo

Applied Mathematics Department
Graduate Program in Telecommunications Engineering
Universidade Federal Fluminense
Rua Mário Santos Braga, S/N, 24020-140, Niterói, Brasil

Christian Soize *

Université Paris-Est , Laboratoire Modélisation et Simulation Multi échelle,
FRE3160 CNRS,
5 bd Descartes, 77454 Marne-la-Vallée, France

Rubens Sampaio

Pontifícia Universidade Católica do Rio de Janeiro
Departamento de Engenharia Mecânica
Rua Marquês de São Vicente 225, 22453-900, Rio de Janeiro, Brazil

Christophe Desceliers

Université Paris-Est , Laboratoire Modélisation et Simulation Multi échelle,
FRE3160 CNRS,
5 bd Descartes, 77454 Marne-la-Vallée, France

Abstract

Preprint submitted to Computational Mechanics *31 March 2012*

This paper is devoted to the construction of a stochastic nonlinear dynamical system for signal generation such as the production of voiced sounds. The dynamical system is highly nonlinear and the output signal generated is very sensitive to a few parameters of the system. In the context of the production of voiced sounds the measurements have a significant variability. We then propose a statistical treatment of the experiments and we developed a probability model of the sensitive parameters in order that the stochastic dynamical system has the capability to predict the experiments in the probability distribution sense. The computational nonlinear dynamical system is presented. The Maximum Entropy Principle is used to construct the probability model. An experimental validation is shown.

Key words: Nonlinear dynamical systems, Signal generation, Uncertainties, Voiced sound production

---

1 Introduction

This paper is devoted to the signal generation by a nonlinear dynamical system modelling the voice production through a mechanical model. Of course, since the voice production system is a very complex system the mechanical model is simplified. Consequently, uncertainties are introduced in the model and must be taken into account. In addition, the introduction of such simplified mechanical model must be validated by experiments. We then propose a stochastic nonlinear dynamical model to produce voiced sounds. Information such as the fundamental frequency of the stochastic output signal is extracted from the model and we present a statistical validation by experiments.

* Corresponding author. Tel: +33 1 60957661, Fax: +33 1 60957799
Email addresses: ecataldo@im.uff.br (Edson Cataldo), christian.soize@univ-paris-est.fr (Christian Soize), rsampaio@puc-rio.br (Rubens Sampaio), christophe.desceliers@univ-paris-est.fr (Christophe Desceliers).

In this paper we use the two-mass model of the vocal folds proposed by Ishizaka & Flanagan (1972), referred henceforth as IF72, which has been widely used and the capability of this well-known model to reproduce the vocal folds vibrations has been successfully demonstrated.

The IF72 model has been used for producing and studying voiced sounds in a deterministic way. The voiced sounds produced by the IF72 model depend on the parameters used in the simulation. However, the parameters are uncertain. To improve the predictions we take into account the uncertainties of the parameters using a parametric probabilistic approach, which consists in modelling each uncertain parameter by a random variable. In principle all the parameters could be taken as uncertain but in this paper we take only some of them. These uncertain parameters are the *neutral glottal area*, the *subglottal pressure* and the *tension parameter*, because the fundamental frequency mainly depends on these parameters.

Since the fundamental frequency is a nonlinear mapping from the random variables modelling the three uncertain parameters, the probability density functions of these random variables are required in order to compute the probability density function of the fundamental frequency. The probability density functions of these three uncertain parameters are then derived using the Maximum Entropy Principle. Such an approach is used because available experimental data sets are not sufficiently large to construct a good estimation of the probability density function using nonparametric mathematical statistics.

Below we are interested in constructing the probability density function of the fundamental frequency of a voice signal (see, for instance, Titze (1994) and, in particular, Pinto & Titze (1990) who discuss distributional characteristics of perturbations and say that different types of vocal perturbations may have different distributions). Since the dynamical system producing the

output stochastic signal is nonlinear, the stochastic solver used is based on the Monte Carlo simulation method. Independent realizations of the random variables modelling the uncertain parameters are obtained with generators adapted to their probability distributions which are constructed in the paper.

The paper is organized as follows. In Sec. 2 the equations that describe the deterministic problem (defined as the mean problem) are formulated; the procedure for computing the output signal and, consequently, the fundamental frequency is presented for the mean problem. In Sec. 3 the construction of the probabilistic model of the uncertain parameters and the construction of the approximation of the stochastic problem are presented. Section 4 is devoted to the calculation of the statistics of the random fundamental frequency and of its probability density function. A comparison of the results with experimental data is in Sec. 5. Conclusions are outlined in Sec. 6. The Appendix describes the functions and matrices used in the model.

2 Equations for the mean model

The mechanical model depends on two sets of parameters. The first one is constituted of fixed parameters for which uncertainties are not taken into account. The second is constituted of three uncertain parameters (introduced in Sec. 1) that will be modeled by random variables in Sec. 3. In this section, all the parameters are considered deterministic and the parameters of the second set are fixed at their mean values.

Fig. 1 is a schematic diagram of the IF72 model, called a source-filter model (Fant, 1960; Ishizaka & Flanagan, 1972). It is made up of two parts: the subsystem of the vocal folds (*source*) and the subsystem of the vocal tract (*filter*). They are coupled by the glottal volume velocity. During phonation, the filter is excited by a sequence of airflow pulses which are periodic signals

with fundamental frequencies equal to the fundamental frequency of the voice signal. Each vocal fold is represented by two (nonlinear) mass-spring-damper systems, coupled through a spring of stiffness k_c. The two vocal folds constitute a symmetric system. The vocal tract is represented by a standard two-tube configuration for vowel /a/ (Ishizaka & Flanagan, 1972; Titze, 1994).

The principal assumptions are that the motion of each vocal fold is perpendicular to the direction of airflow that is assumed to be quasi-steady and is described by Bernoulli's energy equation. The complete model is a well known representation of speech production acoustics. In this paper, the vocal tract is constituted of two coupled tubes with linear acoustics behavior. The subglottal pressure, y, is the input of the *source* subsystem whose output is the air volume velocity, u_g. The air volume velocity is the input of the *filter* subsystem whose output is the radiated pressure, p_r. If we consider the complete system, the subglottal pressure y is the input and the output is the radiated pressure p_r. The main interest of this paper is to analyze the changings of the fundamental frequency of the produced voice signal. The three main parameters of the IF72 model that are responsible for these changings are:

a_{g0}: the area at rest between the vocal folds, called the *neutral glottal area*.
y: the *subglottal pressure*.
q: the *tension parameter* which controls the fundamental frequency of the vocal-fold vibrations .

In order to control the fundamental frequency of the vocal folds, parameters m_1, k_1, m_2, k_2, k_c (see Fig. 1) are written as $m_1 = \widehat{m}_1/q$, $k_1 = q\,\widehat{k}_1$, $m_2 = \widehat{m}_2/q$, $k_2 = q\,\widehat{k}_2$, $k_c = q\,\widehat{k}_c$, in which $\widehat{m}_1$, , $\widehat{k}_1$, $\widehat{m}_2$, $\widehat{k}_2$, $\widehat{k}_c$ are fixed values.

The mean model, proposed by Ishizaka & Flanagan (1972), depends on eight parameters m_1, k_1, m_2, k_2, k_c, a_{g0}, y, q, and is rewritten as:

$$\phi_1(\mathbf{w})|u_g|u_g + \phi_2(\mathbf{w})u_g + \phi_3(\mathbf{w})\dot{u}_g + \frac{1}{\tilde{c}_1}\int_0^t (u_g(\tau) - w_3(\tau))d\tau - y = 0 \quad (1)$$

$$[M]\ddot{\mathbf{w}} + [C]\dot{\mathbf{w}} + [K]\mathbf{w} + \mathbf{h}(\mathbf{w}, \dot{\mathbf{w}}, u_g, \dot{u}_g) = 0 \quad (2)$$

in which $\mathbf{w}(t) = (w_1(t), w_2(t), w_3(t), w_4(t), w_5(t))$, with $w_1(t) = x_1(t)$, $w_2(t) = x_2(t)$, $w_3(t) = u_1(t)$, $w_4(t) = u_2(t)$, $w_5(t) = u_r(t)$. The functions x_1 and x_2 are the displacements of the masses m_1 and m_2, the functions u_1 and u_2 describe the air volume velocity through the two tubes modelling the vocal tract, and u_r is the air volume velocity through the mouth. The function p_r (output signal) is evaluated by $p_r(t) = u_r(t)\tilde{r}_r$, with $\tilde{r}_r = 128\rho v_s/9\pi^3 r_2^2$, where ρ is the air mass density, v_s is the sound velocity, and r_2 is the radius of tube 2. Constant $\tilde{c}_1$, functions ϕ_1, ϕ_2, ϕ_3, $(\mathbf{w}, \dot{\mathbf{w}}, u_g, \dot{u}_g) \mapsto \mathbf{h}(\mathbf{w}, \dot{\mathbf{w}}, u_g, \dot{u}_g)$, and matrices $[M]$, $[C]$, $[K]$ are described in the Appendix.

We can note that Eq.2 describes the vibration problem in each of the two subsystems (vocal folds and vocal tract) and Eq.1 is the equation that couples the two subsystems.

2.1 Solver

In order to solve Eq. (1) and (2), i.e. to find u_g and $\mathbf{w}$ for a given y, an implicit-time numerical method is proposed.

This algorithm uses (1) an implicit *forward finite difference* method for Eq. (1) in which the integral is discretized with the method of *left Riemann sum* and (2) an unconditionally stable *Newmark* method for Eq. (2). Let Δt be the sampling time and $\mathbf{w}_i = \mathbf{w}(i\Delta t)$, $\dot{\mathbf{w}}_i = \dot{\mathbf{w}}(i\Delta t)$, $\ddot{\mathbf{w}}_i = \ddot{\mathbf{w}}(i\Delta t)$, $u_{g_i} = u_g(i\Delta t)$ and $\dot{u}_{g_i} = \dot{u}_g(i\Delta t)$. Then, for all $i \geq 1$, Eqs. (1) and (2) yield

$$\phi_1(\mathbf{w}_i)|u_{g_i}|u_{gi}+\phi_2(\mathbf{w}_i)u_{g_i}+\phi_3(\mathbf{w}_i)\frac{1}{\Delta t}(u_{g_i}-u_{g_{i-1}})+\frac{1}{\tilde{c}_1}\Delta t\sum_{k=0}^{i-1}(u_{g_k}-w_{3_k})-y=0 \tag{3}$$

and

$$[A]\mathbf{w}_i+\mathbf{h}\left(\mathbf{w}_i,\dot{\mathbf{w}}_i,u_{g_i},\frac{u_{g_i}-u_{g_{i-1}}}{\Delta t}\right)=\mathbf{z}_i \tag{4}$$

in which

$$\begin{cases} [A]=[K]+\tilde{a}_0[M]+\tilde{a}_1[C] \\ \mathbf{z}_i=[M](\tilde{a}_0\mathbf{w}_{i-1}+\tilde{a}_2\dot{\mathbf{w}}_{i-1}+\tilde{a}_3\ddot{\mathbf{w}}_{i-1})+[C](\tilde{a}_1\mathbf{w}_{i-1}+\tilde{a}_4\dot{\mathbf{w}}_{i-1}+\tilde{a}_5\ddot{\mathbf{w}}_{i-1}) \\ \ddot{\mathbf{w}}_i=\tilde{a}_0(\mathbf{w}_i-\mathbf{w}_{i-1})-\tilde{a}_2\dot{\mathbf{w}}_{i-1}-\tilde{a}_3\ddot{\mathbf{w}}_{i-1} \\ \dot{\mathbf{w}}_i=\dot{\mathbf{w}}_{i-1}+\tilde{a}_6\ddot{\mathbf{w}}_{i-1}+\tilde{a}_7\ddot{\mathbf{w}}_i \\ \tilde{a}_0=\frac{1}{\tilde{\alpha}\Delta t^2}\,,\ \tilde{a}_1=\frac{\tilde{\delta}}{\tilde{\alpha}\Delta t}\,,\ \tilde{a}_2=\frac{1}{\tilde{\alpha}\Delta t}\,,\ \tilde{a}_3=\tilde{\alpha}-\frac{1}{2} \\ \tilde{a}_4=\frac{\tilde{\delta}}{\tilde{\alpha}}-1\,,\ \tilde{a}_5=\frac{\tilde{\delta}}{2}\left(\frac{\tilde{\delta}}{\tilde{\alpha}}-2\right)\,,\ \tilde{a}_6=\Delta t(1-\tilde{\delta})\,,\ \tilde{a}_7=\tilde{\delta}\Delta t \end{cases} \tag{5}$$

with $u_{g_0}=0$, $\mathbf{w_0}=0$, $\dot{\mathbf{w}}_\mathbf{0}=0$, $\tilde{\delta}=0.5$ and $\tilde{\alpha}=0.25$.

The method used to construct the approximation of Eqs. (3) and (4) consists in finding u_{g_i} as the limit of the sequence $\{u_{g_i}^\alpha\}$, $\alpha\geq 1$, when α goes to infinity such that, for all $\alpha\geq 1$ and $i\geq 1$, we have

$$\phi_1(\mathbf{w}_i^{\alpha-1})|u_{g_i}^\alpha|u_{g_i}^\alpha+\phi_2(\mathbf{w}_i^{\alpha-1})u_{g_i}^\alpha+\phi_3(\mathbf{w}_i^{\alpha-1})\frac{1}{\Delta t}(u_{g_i}^\alpha-u_{g_{i-1}})+\frac{1}{\tilde{c}_1}\Delta t\sum_{k=0}^{i-1}(u_{g_k}-w_{3_k})-y=0\ , \tag{6}$$

with $\mathbf{w}_i^0 = \mathbf{w}_{i-1}$ and $u_{g_i}^1 = u_{g_{i-1}}$. In Eq. (6), $\mathbf{w}_i^{\alpha-1}$ is the limit of the sequence $\{\mathbf{w}_i^{\alpha-1,\beta}\}$, $\beta \geq 0$, when β goes to infinity and is such that, for all $\beta \geq 1$, $\alpha \geq 2$ and $i \geq 1$,

$$[A]\mathbf{w}_i^{\alpha-1,\beta} = \mathbf{z}_i - \mathbf{h}\left(\mathbf{w}_i^{\alpha-1,\beta-1}, \dot{\mathbf{w}}_i^{\alpha-1,\beta-1}, \frac{u_{g_i}^{\alpha-1} - u_{g_{i-1}}}{\Delta t}\right) , \tag{7}$$

with $\mathbf{w}_i^{\alpha,0} = \mathbf{w}_i^{\alpha-1}$ and $\dot{\mathbf{w}}_i^{\alpha,0} = \dot{\mathbf{w}}_i^{\alpha-1}$.

For each time step i, index α of the iteration loop being fixed, first Eq. (6) is solved to calculate u_{gi}^{α} and, second, Eq. (7) is solved to calculate $\mathbf{w}_i^{\alpha-1}$ using an iteration loop in β. Loop in α is performed until convergence is reached. Then, the next time step is computed.

As will be seen from the results, the methodology of approximation was well adapted to the problem.

2.2 Validation of the solver

Numerical tests of the algorithm have been performed and it was verified that it is unconditionally stable. Below we present an example related to the output signal for a vowel /a/ using data from Ishizaka & Flanagan (1972): $d_1 = 2.5 \times 10^{-3}\,\mathrm{m}$, $d_2 = 5 \times 10^{-4}\,\mathrm{m}$, $y = 8000\,\mathrm{Pa}$, $a_{g0} = 5 \times 10^{-6}\,\mathrm{m}^2$, $q = 1$, $m_1 = 1.25 \times 10^{-4}\,\mathrm{kg}$, $m_2 = 2.5 \times 10^{-5}\,\mathrm{kg}$, $k_1 = 80\,\mathrm{N/m}$, $k_2 = 8\,\mathrm{N/m}$, $k_c = 25\,\mathrm{N/m}$, $\xi_1 = 0.1$, $\xi_2 = 0.6$, $\eta_{k_1} = \eta_{k_2} = 100$, $\eta_{h_1} = \eta_{h_2} = 500$. The lengths considered for the two tubes are $\ell_1 = 8.9 \times 10^{-2}\,\mathrm{m}$ and $\ell_1 = 8.1 \times 10^{-2}\,\mathrm{m}$, and their corresponding radius are $r_1 = 0.56 \times 10^{-2}\,\mathrm{m}$ and $r_2 = 1.49 \times 10^{-2}\,\mathrm{m}$. The sampling time step used to set a good accuracy is $\Delta t = 1/45000\,\mathrm{s}$. Figure 2 displays the time response of the system for u_g, x_1, x_2, and p_r normalized to $p_{max} = \max_t p_r(t)$. The graphs agree with those published by Ishizaka & Flanagan (1972).

2.3 Calculation of the output signal fundamental frequency

As we have explained in Sec. 1, the objectives of this paper are: (1) to make a probabilistic analysis of the fundamental frequency f_0 of the output signal p_r and (2) to construct the probability density function of f_0. Let t_{max} be such that $p_r(t) = 0$, for $t \geq t_{max}$. The Fourier transform of $t \mapsto p_r(t)$, denoted by $\omega \mapsto \widehat{p}_r(\omega)$, is given by

$$\widehat{p}_r(\omega) = \int_T p_r(t) e^{-i\omega t}\, dt \ , \tag{8}$$

in which $T = [0, t_{max}]$. The fundamental frequency is then defined as the frequency of the first peak in the graph of $\omega \mapsto \mid \widehat{p}_r(\omega) \mid$. Clearly, there are two mappings $\mathcal{L}$ and $\mathcal{M}$ such that the output signal p_r at time t and the fundamental frequency f_0 can be written as

$$p_r(t) = \mathcal{L}(t; a_{g0}, y, q) \ , \tag{9}$$

$$f_0 = \mathcal{M}(a_{g0}, y, q) \, . \tag{10}$$

For instance, Fig. 3 shows the modulus of the Fourier transform $|\widehat{p}_r(\omega)|$ normalized with respect to $\widehat{p}_{max} = \max_\omega \mid \widehat{p}_r(\omega) \mid$ associated with the time signal shown in Fig. 2, in which the first peak is marked. If an output signal is not produced, then f_0 is taken as zero.

3 Stochastic modelling

The three main parameters responsible for the changing of the fundamental frequency will be considered as uncertain and random variables will be as-

sociated with them. It means that for each realization of the three random variables a different output signal is produced, that is the output signal is a stochastic process. It is assumed that the stochastic process can locally be modelled as being stationary and ergodic stochastic process (see, for instance, Schoengten (2001)).

The probability density functions associated with the random variables corresponding to the chosen uncertain parameters will be constructed using the Maximum Entropy Principle (see Jaynes's (1957a,1857b)) in the context of the Information theory introduced by Shannon (1948).

This principle states: *Out of all probability distributions consistent with a given set of available information, choose the one that has maximum uncertainty (entropy).*

The measure of uncertainty (entropy) used is given by Eq.11:

$$S(p_X) = -\int_{-\infty}^{+\infty} p_X(x) ln\,(\,p_X(x)\,)\,dx\,. \tag{11}$$

The goal is to maximize the entropy S, under the constraints defined by the following available information

$$\int_{-\infty}^{+\infty} p_X(x)dx = 1 \;\text{ and }\; \int_{-\infty}^{+\infty} p_X(x) g_i(x)dx = a_i\;,\; i = 1, \ldots, m \tag{12}$$

where the real numbers a_i and the functions g_i are given.

3.1 Probabilistic model of the uncertain parameters

As explained in Sec. 1, the three parameters a_{g0}, y, and q are modeled by random variables A_{g0}, Y, and Q. Consequently, parameters m_1, k_1, m_2, k_2,

and k_c become random variables denoted by M_1, K_1, M_2, K_2, and K_c defined by $M_1 = \widehat{m}_1/Q$, $K_1 = Q\widehat{k}_1$, $M_2 = \widehat{m}_2/Q$, $K_2 = Q\widehat{k}_2$, and $K_c = Q\widehat{k}_c$. The probability models derived here are particular cases of those ones described in Soize (2001). Since no information is available concerning cross statistical informations between random variables A_{g0}, Y, Q, the use of the Maximum Entropy Principle shows that these random variables are independent. The level of uncertainties will be controlled by the coefficients of variation $\delta_{A_{g0}}$, δ_Y and δ_Q of the random variables A_{g0}, Y and Q and will de defined as dispersion parameters of the probability model.

3.1.1 Random variable A_{g0}

The parameter a_{g0} is modelled by a random variable A_{g0} with values in $\mathbb{R}^+$ (due to physical restriction) and the mean value $\underline{A}_{g0}$ is known. Then the available information can be defined as (1) the support of the probability density function which is $]0, +\infty[$, (2) the mean value which is such that $E\{A_{g0}\} = \underline{A}_{g0}$, (3) the second-order moment which must be finite, $E\{A_{g0}^2\} < +\infty$. The probability density function $p_{A_{g0}}$ of A_{g0} has then to verify the following constraint equations,

$$\int_{-\infty}^{+\infty} p_{A_{g0}}(a_{g0})\, da_{g0} = 1\,,\ \int_{-\infty}^{+\infty} a_{g0}\, p_{A_{g0}}(a_{g0})\, da_{g0} = \underline{A}_{g0}\ ,\ \int_{-\infty}^{+\infty} a_{g0}^2\, p_{A_{g0}}(a_{g0})\, da_{g0} = c, \tag{13}$$

in which c is a positive finite constant which is unknown. The use of the Maximum Entropy Principle yields:

$$p_{A_{g0}}(a_{g0}) = \mathbf{1}_{]\mathbf{0},+\infty[}\, e^{-\lambda_0 - \lambda_1 a_{g_0} - \lambda_2 (a_{g_0})^2}, \tag{14}$$

where λ_0, λ_1 and λ_2 are the solution of the three equations defined by Eq. 13.

Since the constant c is unknown, we introduce a new parametrization expressing c as a function of the coefficient of variation $\delta_{A_{g0}}$ of the random variable A_{g0} which is such that $\delta_{A^2_{g0}} = E\{A^2_{g0}\}/\underline{A}^2_{g0} - 1$ which implies that $c = \underline{A}_{g0}{}^2 \left(1 + \delta^2_{A_{g0}}\right)$.

3.1.2 Random variable Y

The parameter y is modelled by a random variable Y with values in $\mathbb{R}^+$ (due to physical restriction) and the mean value $\underline{Y}$ is known. Then the available information is constituted of (1) the support of the probability density function which is $]0, +\infty[$, (2) the mean value which is such that $E\{Y\} = \underline{Y}$, (3) the condition $E\{ln(Y)\} = c_1$ with $|c_1| < +\infty$ which implies that zero is a repulsive value for the positive-valued random variable Y. The introduction of the last available information is related to the need to have a minimum value of Y to produce an output signal (Baer, 1975). The probability density function p_Y of Y has then to verify the following constraint equations:

$$\int_{-\infty}^{+\infty} p_Y(y)\, dy = 1 \quad , \quad \int_{-\infty}^{+\infty} y\, p_Y(y)\, dy = \underline{Y} \quad , \quad \int_{-\infty}^{+\infty} ln(y)\, p_Y(y)\, dy = c_1 \,. \tag{15}$$

Applying the Maximum Entropy Principle yields the following probability density function for Y,

$$p_Y(y) = \mathbf{1}_{]0,+\infty[}(y) \frac{1}{\underline{Y}} \left(\frac{1}{\delta^2_Y}\right)^{\frac{1}{\delta^2_Y}} \frac{1}{\Gamma\left(1/\delta^2_Y\right)} \left(\frac{y}{\underline{Y}}\right)^{\frac{1}{\delta^2_Y}-1} exp\left(-\frac{y}{\delta^2_Y \underline{Y}}\right) , \tag{16}$$

in which $\delta_Y = \sigma_Y/\underline{Y}$ is the coefficient of variation of the random variable Y such that $0 \leq \delta_Y < 1/\sqrt{2}$ and where σ_Y is the standard deviation of Y. In this equation $\alpha \mapsto \Gamma(\alpha)$ is the Gamma function defined by $\Gamma(\alpha) = \int_0^{+\infty} t^{\alpha-1} e^{-t} dt$.

From Eq. (16), it can be verified that Y is a second-order random variable.

3.1.3 Random variable Q

The parameter q is modelled by a random variable Q with values in $\mathbb{R}^+$ (due to physical restriction) and the mean value $\underline{Q}$ is known. Since $M_1 = \widehat{m}_1/Q$ has to be a second-order random variable, it is necessary that $E\{M_1^2\} < +\infty$ yielding $E\{1/Q^2\} < +\infty$. Then the available information is defined by (1) the support of the probability density function is $]0, +\infty[$, (2) the mean value is such that $E\{Q\} = \underline{Q}$ and (3) $E\{1/Q^2\} = c_2'$ with $c_2' < +\infty$. The third constraint is taken into account by requiring that $E\{ln(Q)\} = c_2$ with $|c_2| < +\infty$. So, the probability density function p_Q of Q, whose support is $]0, +\infty[$, has to verify the following constraint equations,

$$\int_{-\infty}^{+\infty} p_Q(q)\, dq = 1 \quad , \quad \int_{-\infty}^{+\infty} q\, p_Q(q)\, dq = \underline{Q} \quad , \quad \int_{-\infty}^{+\infty} ln(q)\, p_Q(q)\, dq = c_2\, . \tag{17}$$

Applying the Maximum Entropy Principle yields again

$$p_Q(q) = \mathbf{1}_{]0,+\infty[}(q) \frac{1}{\underline{Q}} \left(\frac{1}{\delta_Q^2}\right)^{\frac{1}{\delta_Q^2}} \frac{1}{\Gamma\left(1/\delta_Q^2\right)} \left(\frac{q}{\underline{Q}}\right)^{\frac{1}{\delta_Q^2}-1} exp\left(-\frac{q}{\delta_Q^2 \underline{Q}}\right) \ , \tag{18}$$

where the positive parameter $\delta_Q = \sigma_Q/\underline{Q}$ is the coefficient of variation of the random variable Q such that $\delta_Q < 1/\sqrt{2}$ and where σ_Q is the standard deviation of Q. From Eq. (18), it can be verified that Q is a second-order random variable and that $E\{1/Q^2\} < +\infty$.

3.2 Uncertain mechanical system

The stochastic system is deduced from the deterministic one substituting a_{g0}, y and q by the random variables A_{g0}, Y and Q. Consequently, according to Eq. (10), the random fundamental frequency F_0 is given by $F_0 = \mathcal{M}(A_{g0}, Y, Q)$. However, the nonlinear mapping $\mathcal{M}$ is not explicitly known and it is implicitly defined by Eqs. (1), (2), (8), and (10) substituting a_{g0}, y and q by random variables A_{g0}, Y and Q.

3.3 Stochastic solver for the uncertain mechanical system

Equations (1), (2), (8), (10) defining the nonlinear mapping $\mathcal{M}$ have to be solved using their approximations defined by Eqs. (3) to (7) and (8) substituting a_{g0}, y and q by the random variables A_{g0}, Y and Q. The stochastic solver used is based on the Monte Carlo method. First, independent realizations $\mathbf{X}(\theta)$ of the random variable $\mathbf{X} = (A_{g0}, Y, Q)$ are constructed using the probability density functions defined by Eqs. (14), (16) and (18). For each realization $\mathbf{X}(\theta)$, the realization $F_0(\theta)$ of the random fundamental frequency F_0 is given by

$$F_0(\theta) = \mathcal{M}(A_{g0}(\theta), Y(\theta), Q(\theta))\,, \tag{19}$$

and is calculated solving deterministic Eqs. (3) to (7) and (8) substituting $x = (a_{g0}, y, q)$ by $X(\theta) = (A_{g0}(\theta), Y(\theta), Q(\theta))$. The mean-square convergence of the random variable F_0 is analyzed with respect to the number n of independent realizations for the Monte Carlo method. The mathematical statistics are used to construct the estimation of (1) the mean value $m_{F_0} = E\{F_0\}$, (2) the variance $\sigma^2_{F_0}$ of the random variable F_0, (3) the confidence region of the random variable F_0 and, finally, (4) the probability density function p_{F_0}.

3.3.1 Mean-square convergence analysis

The mean-square convergence analysis with respect to independent realizations $F_0(\theta_1), \ldots, F_0(\theta_n)$ of the random variable F_0 is carried out studying the function $n \mapsto \mathrm{Conv}(n)$ defined by

$$\mathrm{Conv}(n) = \frac{1}{n}\sum_{j=1}^{n} F_0(\theta_j)^2. \tag{20}$$

This convergence analysis is performed for different values of $\delta_{A_{g0}}$, δ_Y, and δ_Q. For $n \geq 2000$, the convergence is always reached. Then, $n = 2000$ was used for all further estimations. For instance, Fig. 4 shows the graph of the function $n \to log_{10}(\mathrm{Conv}(n))$ for $\delta_{A_{g0}} = \delta_Y = \delta_Q = 0.10$.

3.3.2 Estimation of the mean value, variance, confidence region, and probability density function of F_0

An estimation $\widehat{m}_{F_0}$ of the mean value $m_{F_0} = E\{F_0\}$ and an estimation $\widehat{\sigma_{F_0}}^2$ of the variance $\sigma^2_{F_0}$ of the random variable F_0 are given by

$$\widehat{m}_{F_0} = \frac{1}{n}\sum_{j=1}^{n} F_0(\theta_j) \ , \tag{21}$$

$$\widehat{\sigma}^2_{F_0} = \frac{1}{n-1}\sum_{j=1}^{n} \left(F_0(\theta_j) - \widehat{m}_{F_0}\right)^2 \ . \tag{22}$$

The confidence region associated with a probability level P_c is constructed using quantiles (Serfling, 1980). Let $F_{F_0}(f_0) = P\{F_0 \leq f_0\}$ be the cumulative distribution function of the random variable F_0. For $0 < p < 1$, the $p-$th quantile of F_{F_0} is defined as $\zeta(p) = \mathrm{Inf}\{f : F_{F_0}(f) \geq p\}$. Then, the upper envelope f^+ and the lower envelope f^- of the confidence interval are defined by $f^+ = \zeta((1+P_c)/2)$ and $f^- = \zeta((1-P_c)/2)$. Let $f_1 = F_0(\theta_1), \ldots, f_n = F_0(\theta_n)$

be n independent realizations of the random variable F_0. Let $\widetilde{f}_1 < \ldots < \widetilde{f}_n$ be the order statistics associated with $f_1, \ldots, f_n$. Therefore, we have the following estimation: $f^+ = \widetilde{f}_{j^+}$ with $j^+ = \text{fix}(n(1+P_c)/2)$ and $f^- = \widetilde{f}_{j^-}$ with $j^- = \text{fix}(n(1-P_c)/2)$ in which $\text{fix}(z)$ is the integer part of the real number z.

The estimation of the probability density function p_{F_0} of the random variable F_0 is constructed as follows. Let M be the number of intervals. Let $I_j = [\nu_j, \nu_j + \Delta\nu[$ for $j = 1, \ldots, M$ with $\nu_1 = \widetilde{f}_1$ and $\Delta\nu = (\widetilde{f}_M - \widetilde{f}_1)/M$. An estimation $\widehat{p}_{F_0}$ of the probability density function of F_0 is given by

$$\widehat{p}_{F_0}(f_0) = \sum_{j=1}^{M} \mathbf{1}_{I_j}(f_0) \frac{n_j}{n\Delta\nu} \,. \tag{23}$$

in which n_j is the number of realizations in the interval I_j.

4 Application to the production of voiced sounds

The application to the production of voiced sounds is done with the data defined in Sec. 2.2. The mean values of the uncertain parameters used are $\underline{A}_{g0} = 5 \times 10^{-2}$ m^2, $\underline{Y} = 800$ Pa, and $\underline{Q} = 1$. It should be noted that the fundamental frequency f_0 calculated with the deterministic model ($a_{g0} = \underline{A}_{g0}$, $y = \underline{Y}$, and $q = \underline{Q}$) is $f_0 = 168.5$ Hz.

Figure 5 shows the confidence regions for the fundamental frequency F_0 with (1) $\delta_Y = 0.01$, δ_Q varying from 0.01 up to 0.60 (top figure); (2) $\delta_Q = 0.01$, δ_Y varying from 0.01 up to 0.60 (bottom figure). For both cases, a_{g0} was considered fixed at $0.05\,\text{cm}^2$. In each figure, the middle line is the estimation $\widehat{m}_{F_0}$ of the mean value defined by Eq. (21). Figure 6 displays the probability density function of F_0 for five different cases with respect to different values of $\delta_{A_{g0}}$, δ_Y, δ_Q defined in the figures. For a realization θ, when no voiced sound is produced, the realization $F_0(\theta)$ of the random fundamental frequency F_0 is

taken as zero.

Figure 5 shows that the dispersion of the random fundamental frequency and also the values of $\widehat{m}_{F_0}$ increase with the level of uncertainties. However, it should be noted that this dispersion is much more important with respect to δ_Q than with respect to δ_Y. In addition, the probabilistic approach which is proposed allows a quantification of uncertainties propagation through the model to be performed. Such a quantification can be analyzed constructing the probability density function of F_0 shown in Fig.6.

It should be noted that the probabilistic approach of uncertainties is particularly well adapted to characterize (or to identify) the probabilistic model of voice signals produced by a given person taking into account the dispersion.

5 Experimental validation

In order to validate the development presented here experimental voice signals produced by one person have been analyzed and their statistics have been compared with simulations with the mechanical model with uncertainties developed here. The measurements are made up of 675 recorded voice signals corresponding to a sustained vowel /a/ from one person. The duration of each signal is 0.01 s. For each experimental signal the corresponding experimental fundamental frequency is calculated. The mean value of the experimental fundamental frequency is $m_{F_0} = 120.77$ Hz and its coefficient of variation is $\delta_{F_0} = 0.0173$. In addition, the experimental probability density function is calculated. The objective is to compare the probability density function of the experimental fundamental frequency with the probability density function constructed with the stochastic mechanical model. The methodology used is the following:

Step 1: The values of a_{g_0}, y and q are identified to obtain the experimental value $f_0 = 120.77$ Hz of the fundamental frequency with the deterministic model.

Step 2: The values of a_{g_0}, y and q found in Step 1 are used as the mean values $\underline{A}_{g_0}$, $\underline{Y}$ and $\underline{Q}$ of the random variables A_{g_0}, Y and Q .

Step 3: With the mean values defined in Step 2, the values of $\delta_{A_{g0}}$, δ_Y and δ_Q are identified to obtain the experimental value $\delta_{F_0} = 0.0173$ of the coefficient of variation of the fundamental frequency with the stochastic mechanical model.

The numerical results obtained using this methodology are the followings:

Step 1: $a_{g0} = 5 \times 10^{-2}\,\mathrm{m}^2$, $y = 750\,\mathrm{Pa}$ and $q = 0.66$.

Step 2: The mean values are $\underline{A}_{g0} = 5 \times 10^{-2}\,\mathrm{m}^2$, $\underline{Y} = 750\,\mathrm{Pa}$ and $\underline{Q} = 0.66$.

Step 3: With the mean values described in Step 2, the mean value of the fundamental frequency obtained, considering 700 realizations in the Monte Carlo method is $m_{F_0} = 120.95\,\mathrm{Hz}$. With the values of the dispersion parameters $\delta_{A_{g0}} = 0.03$, $\delta_Y = 0.01$ and $\delta_Q = 0.01$, the value obtained for the coefficient of variation of the fundamental frequency is $\delta_{F_0} = 0.0171$ which has to be compared with the experimental value 0.0173.

Figure 7 shows the comparison of the probability density function constructed from experimental signals (top) with the probability density function constructed from simulations (bottom). The figures show a reasonably good agreement.

6 Conclusions

We have proposed a parametric probabilistic approach to take into account uncertainties in a nonlinear dynamical model used to produce voiced sounds.

The three parameters controling the values of fundamental frequency of the produced voice signal are modelled by random variables whose probability distributions are constructed using the Maximum Entropy Principle. A complete stochastic computational model has been developed. An experimental validation is presented in the statistical sense.

In the context of such a problem, the experimental output signal of the nonlinear dynamical system has a significant variability which has to be analyzed in the context of the probability theory. We have shown that it is possible to identify a reasonable computational stochastic nonlinear dynamical model which allows experiments to be predicted in the probability distribution sense. This kind of problem is not trivial due to the high nonlinearities in the dynamical system for which a pertinent probability model must be constructed.

Appendix

This appendix defines functions and matrices introduced in Sec. 2.

$$[M] = \begin{bmatrix} m_1 & 0 & 0 & 0 & 0 \\ 0 & m_2 & 0 & 0 & 0 \\ 0 & 0 & \tilde{\ell}_1 + \tilde{\ell}_2 & 0 & 0 \\ 0 & 0 & 0 & \tilde{\ell}_2 + \tilde{\ell}_r & -\tilde{\ell}_r \\ 0 & 0 & 0 & -\tilde{\ell}_r & \tilde{\ell}_r \end{bmatrix} \quad , \quad [C] = \begin{bmatrix} c_1 & 0 & 0 & 0 & 0 \\ 0 & c_2 & 0 & 0 & 0 \\ 0 & 0 & \tilde{r}_1 + \tilde{r}_2 & 0 & 0 \\ 0 & 0 & 0 & \tilde{r}_2 & 0 \\ 0 & 0 & 0 & 0 & \tilde{r}_r \end{bmatrix}$$

$$[K] = \begin{bmatrix} k_1 + k_c & -k_c & 0 & 0 & 0 \\ -k_c & k_2 + k_c & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\tilde{c}_1} + \frac{1}{\tilde{c}_2} & -\frac{1}{\tilde{c}_2} & 0 \\ 0 & 0 & -\frac{1}{\tilde{c}_2} & \frac{1}{\tilde{c}_2} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{h}(\mathbf{w}, \dot{\mathbf{w}}, u_g, \dot{u}_g) = \begin{bmatrix} s_1(w_1) + t_1(w_1)\dot{w}_1 - f_1(w_1, u_g, \dot{u}_g) \\ s_2(w_2) + t_2(w_2)\dot{w}_2 - f_2(w_1, w_2, u_g, \dot{u}_g) \\ -\frac{1}{\tilde{c}_1}u_g \\ 0 \\ 0 \end{bmatrix},$$

with

$\tilde{\ell}_n = \frac{\rho \ell_n}{2\pi r_n^2}$, $\tilde{\ell}_r = \frac{8\rho}{3\pi^2 r_n}$, $\tilde{r}_n = \frac{2}{r_n}\sqrt{\rho\mu\frac{\omega}{2}}$, $\omega = \sqrt{\frac{k_1}{m_1}}$, $a_n = \pi r_n^2$, $\tilde{c}_n = \frac{\ell_n \pi r_n^2}{\rho v_c^2}$, and where ℓ_n is the length of the $n-$th tube, r_n is the radius of the $n-$th tube, and μ is the shear viscosity coefficient.

$\phi_1(\mathbf{w}) = (\frac{0.19\rho}{a_{g0}+2\ell_g w_1} + 2\ell_g w_1) + \frac{\rho}{(a_{g0}+2\ell_g w_2)^2}\left[0.5 - \frac{a_{g0}+2\ell_g w_2}{a_1}\left(1 - \frac{a_{g0}+2\ell_g w_2}{a_1}\right)\right]$
$\phi_2(\mathbf{w}) = (12\mu\ell_g \frac{d_1}{(a_{g0}+2\ell_g w_1)^3} + 12\ell_g^2 \frac{d_2}{(a_{g0}+2\ell_g w_2)^3} + \tilde{r}_1)$, $\phi_3(\mathbf{w}) = (\frac{\rho d_1}{a_{g0}+2\ell_g w_1} + \frac{\rho d_2}{a_{g0}+2\ell_g w_2} + \tilde{\ell}_1)$

$$s_1(w_1) = \begin{cases} k_1\eta_{k_1} w_1^3 \ , \ w_1 > -\frac{a_{g0}}{2\ell_g} \\ k_1\eta_{k_1} w_1^3 + 3k_1\left\{\left(w_1 + \frac{a_{g0}}{2\ell_g}\right) + \eta_{h_1}\left(w_1 + \frac{a_{g0}}{2\ell_g}\right)^3\right\} \ , \ w_1 \leq -\frac{a_{g0}}{2\ell_g} \end{cases}$$

$$s_2(w_2) = \begin{cases} k_2\eta_{k_2}w_2^3\ ,\ w_2 > -\frac{a_{g0}}{2\ell_g} \\ k_2\eta_{k_2}w_2^3 + 3k_2\left\{\left(w_2 + \frac{a_{g0}}{2\ell_g}\right) + \eta_{h_2}\left(w_2 + \frac{a_{g0}}{2\ell_g}\right)^3\right\}\ ,\ w_2 \leq -\frac{a_{g0}}{2\ell_g} \end{cases}$$

$$t_1(w_1) = \begin{cases} 0\,,\ w_1 > -\frac{a_{g0}}{2\ell_g} \\ 2\xi\sqrt{m_1k_1}\,,\ w_1 \leq -\frac{a_{g0}}{2\ell_g} \end{cases} \quad , \quad t_2(w_2) = \begin{cases} 0\,,\ w_2 > -\frac{a_{g0}}{2\ell_g} \\ 2\xi\sqrt{m_2k_2}\,,\ w_2 \leq -\frac{a_{g0}}{2\ell_g} \end{cases}$$

$$f_1(w_1, u_g, \dot{u}_g) = \begin{cases} \ell_g d_1 p_{m_1}(w_1, u_g, \dot{u}_g)\,,\ w_1 > -\frac{a_{g0}}{2\ell_g} \\ 0\,,\ otherwise \end{cases}$$

$$p_{m_1}(w_1, u_g, \dot{u}_g) = y - 1.37\frac{\rho}{2}\left(\frac{u_g}{a_{g0}+2\ell_g w_1}\right)^2 - \frac{1}{2}\left(12\mu\ell_g\frac{d_1}{(a_{g0}+2\ell_g w_1)^3} + \frac{\rho d_1}{a_{g0}+2\ell_g w_1}\right)\dot{u}_g$$

$$f_2(w_1, w_2, u_g, \dot{u}_g) = \begin{cases} \ell_g d_2 p_{m_2}(w_1, w_2, u_g, \dot{u}_g)\,,\ w_1 > -\frac{A_{g0}}{2\ell_g}\ \text{and}\ w_2 > -\frac{a_{g0}}{2\ell_g} \\ \ell_g d_2 y\,,\ w_1 > -\frac{a_{g0}}{2\ell_g}\ \text{and}\ w_2 \leq -\frac{a_{g0}}{2\ell_g} \\ 0\ ,\ \text{otherwise} \end{cases}$$

$$p_{m_2}(w_1, w_2, u_g, \dot{u}_g) = p_{m_1}$$
$$- = \frac{1}{2}\left\{(12\mu\ell_g\frac{d_1}{(a_{g0}+2\ell_g w_1)^3} + 12\ell_g^2\frac{d_2}{(a_{g0}+2\ell_g w_2)^3})u_g + (\frac{\rho d_1}{a_{g0}+2\ell_g w_1} + \frac{\rho d_2}{a_{g0}+2\ell_g w_2})\dot{u}_g\right\}$$
$$-\frac{\rho}{2}u_g^2\left(\frac{1}{(a_{g0}+2\ell_g w_2)^2} - \frac{1}{(a_{g0}+2\ell_g w_1)^2}\right)$$

Acknowledgements

This work was financed by the International Cooperation Project CAPES-COFECUB N.476/04, by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) and by FAPERJ (Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro).

References

Baer, T., Investigation of phonation using excised larynges, Ph.D. thesis, Massachussetts Institute of Technology (Cambridge, MA, 1975).

Fant, G., The acoustic theory of speech production (Mouton, The Hague, 1960).

Ishizaka, K., & Flanagan, J. L., Synthesis of voiced sounds from a two-mass model of the vocal cords, 1972, Bell Systems Tech. J., 51, 1233-1268.

Jaynes, E., Information theory and statistical mechanics, 1957a, Physics Review, 106, 4, 620-630.

Jaynes, E., Information theory and statistical mechanics II, 1957b, Physics Review, 108, 171-190.

Pinto, N., & Titze, I., Unification of pertubation measures in speech signals, 1990, Journal of the Acoustical Society of America, 87 (3), 1278-1289.

Schoengten, J., Stochastic models of jitter, 2001, Journal of the Acoustical Society of America, 109, 1631-1650.

Serfling, R., Approximation theorems of mathematical statistics (Wiley, New York, 1980).

Shannon, C.E., A mathematical theory of communication, 1948, Bell System Tech. Journal, 27, 379-423, 623-659.

Soize, C., Maximum entropy approach for modelling random uncertainties in transient elastodynamics, 2001, Journal of the Acoustical Society of America, 109, 1979-1996.

Titze, I. R., Principles of voice production (Prentice-Hall, Englewood Cliffs,

NJ, 1994).

Captions Accompanying Each Figure

Fig. 1. IF72 model scheme.

Fig. 2. (Color online) Simulation of a vowel /a/: Glottal volume velocity (u_g) (top), displacements of the two masses (x_1 and x_2) (middle), and output signal normalized (p_r) (bottom).

Fig. 3. (Color online) Normalized modulus of the Fourier transform of the output signal. The marker indicates the first peak that corresponds to the fundamental frequency (in this case 160 Hz).

Fig. 4. (Color online) Mean-square convergence of the Monte Carlo method with respect to the number n of realizations.

Fig. 5. (Color online) Confidence regions and mean values of the random fundamental frequency F_0 versus the dispersion coefficient: $\delta_Y = 0.01$ and δ_Q varies (top figure); $\delta_Q = 0.01$ and δ_Y varies (bottom figure). The value of $\underline{A}_{g0}$ is maintained fixed.

Fig. 6. (Color online) Probability density functions of the random fundamental frequency F_0 for five different cases of dispersion parameters.

Fig. 7.(Color online) First case: Experimental probability density function of the fundamental frequency (top) compared with the probability density function estimated with the stochastic model (bottom).

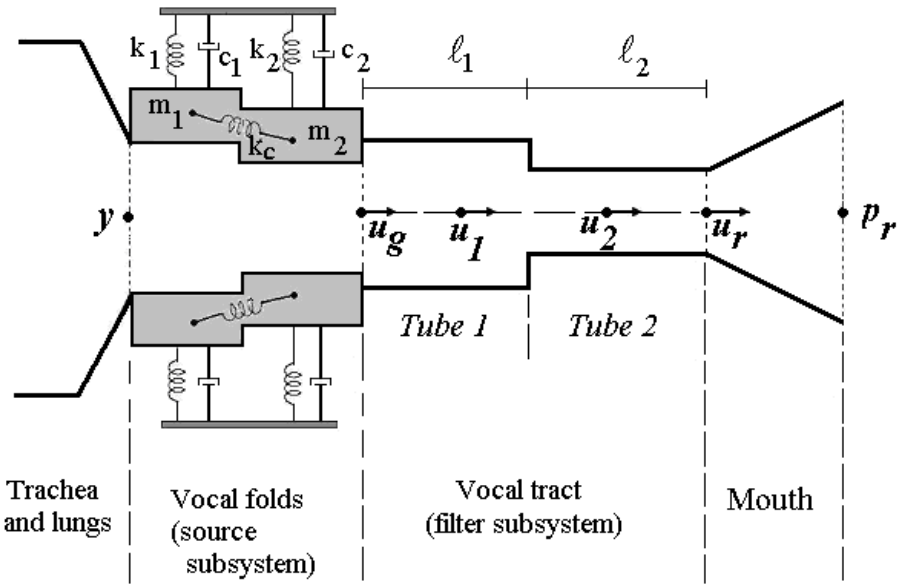
k1
c1
k2
c2
ℓ1
ℓ2
m1
kc
m2
y
ug
u1
u2
ur
pr
Tube 1
Tube 2
Trachea and lungs
Vocal folds (source subsystem)
Vocal tract (filter subsystem)
Mouth

Fig. 1.

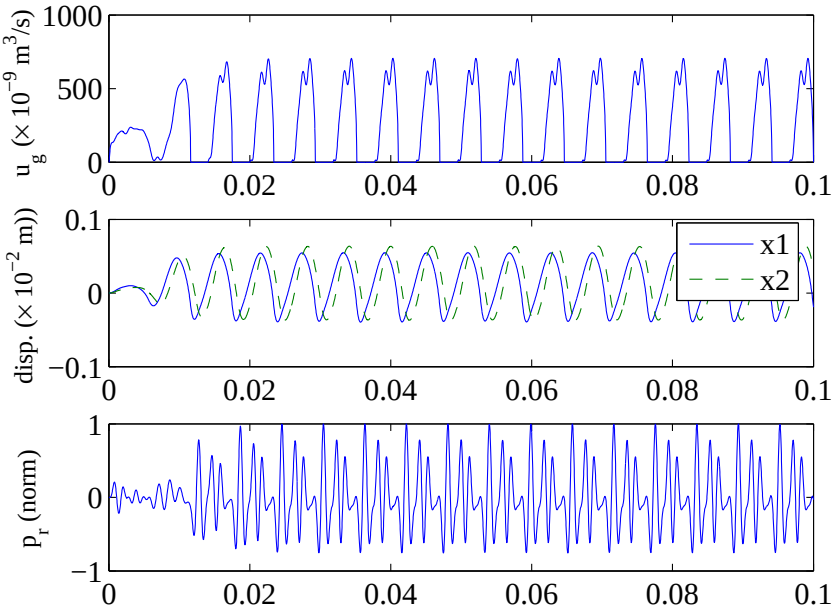
1000
500
0
ug (× 10−9 m3/s)
0
0.02
0.04
0.06
0.08
0.1
0.1
0
−0.1
disp. (× 10−2 m))
x1
x2
1
0
−1
pr (norm)

Fig. 2.

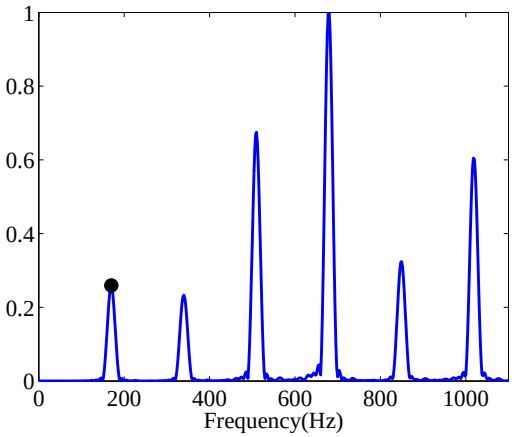
1
0.8
0.6
0.4
0.2
0
0
200
400
600
800
1000
Frequency(Hz)

Fig. 3.

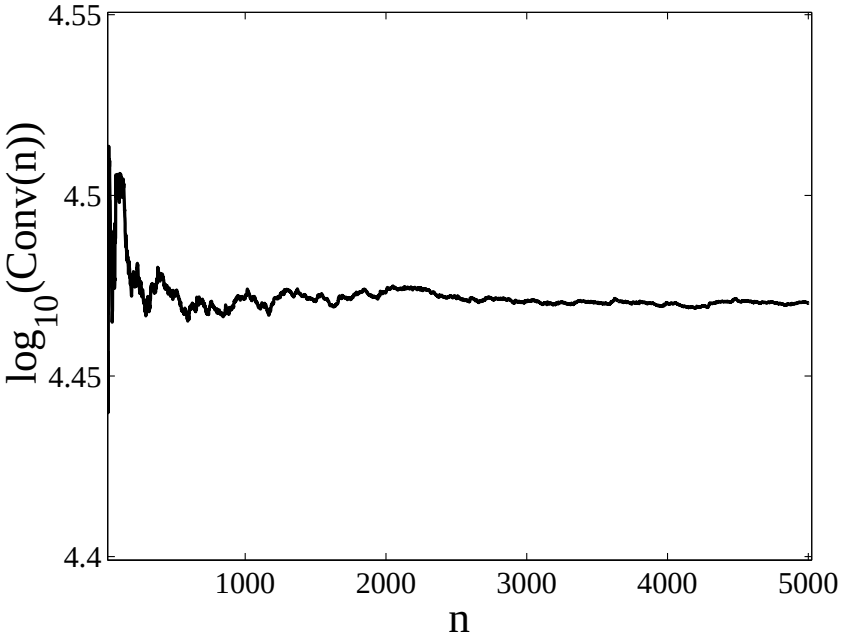
4.55
4.5
4.45
4.4
log10(Conv(n))
1000
2000
3000
4000
5000
n

Fig. 4.

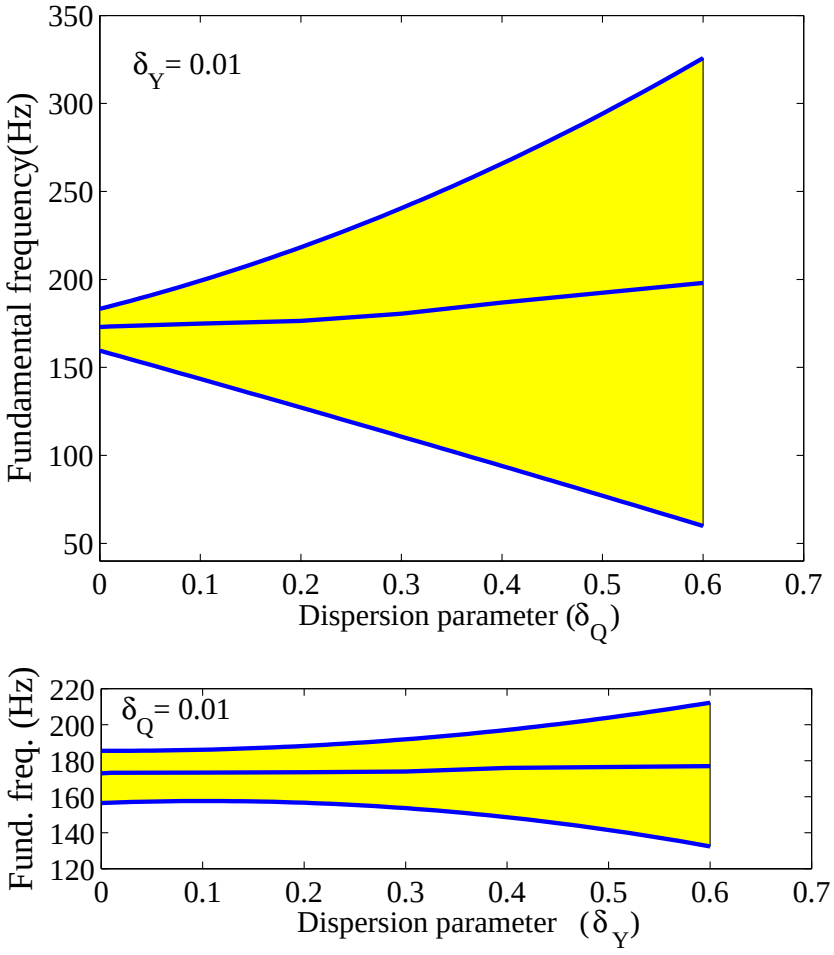
350
300
250
200
150
100
50
Fundamental frequency(Hz)
δY= 0.01
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Dispersion parameter (δQ)
220
200
180
160
140
120
Fund. freq. (Hz)
δQ= 0.01
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Dispersion parameter (δY)

Fig. 5.

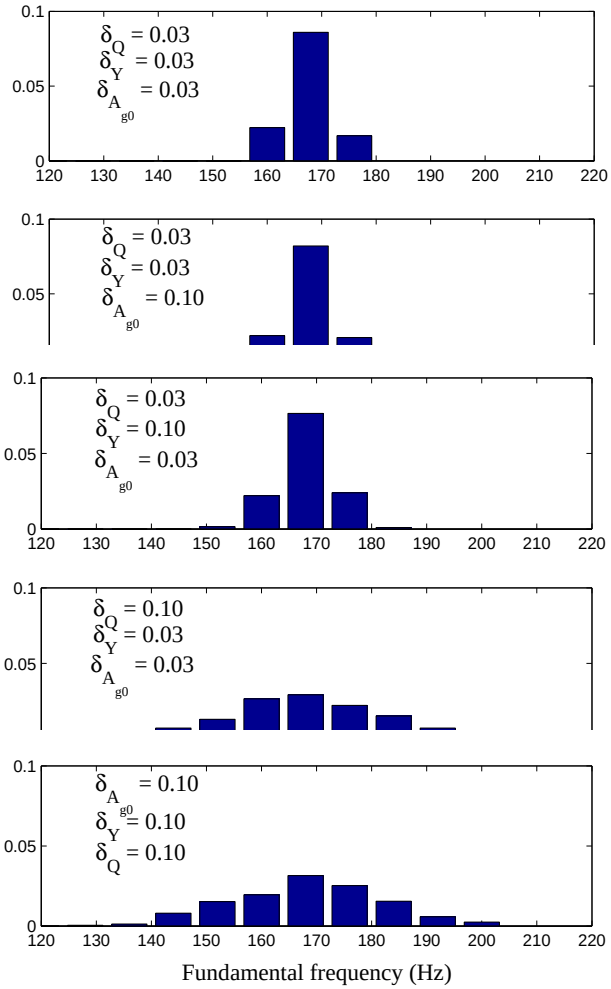
δQ = 0.03
δY = 0.03
δAg0 = 0.03
δQ = 0.03
δY = 0.03
δAg0 = 0.10
δQ = 0.03
δY = 0.10
δAg0 = 0.03
δQ = 0.10
δY = 0.03
δAg0 = 0.03
δAg0 = 0.10
δY = 0.10
δQ = 0.10
0.1
0.05
0
120 130 140 150 160 170 180 190 200 210 220
Fundamental frequency (Hz)

Fig. 6.

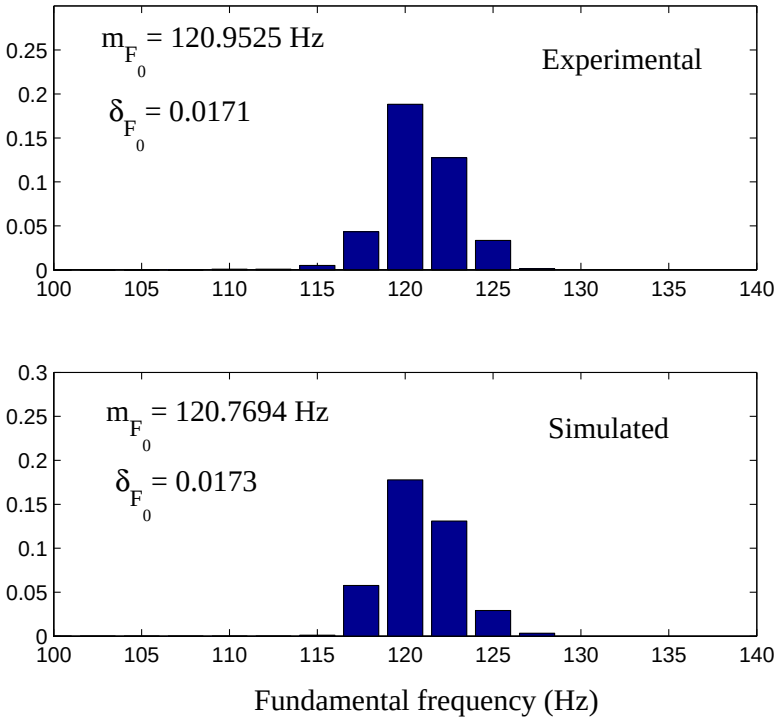
m F0 = 120.9525 Hz
δ F0 = 0.0171
Experimental
m F0 = 120.7694 Hz
δ F0 = 0.0173
Simulated
0.3
0.25
0.2
0.15
0.1
0.05
0
100
105
110
115
120
125
130
135
140
Fundamental frequency (Hz)

Fig. 7.