Prior and posterior probabilistic models of uncertainties in a model for producing voice

Edson Cataldo, Rubens Sampaio, Christian Soize

Universidade Federal Fluminense, Applied Mathematics Departament and Graduate program in Telecommunications Engineering, Rua Mário Santos Braga, S/N, Centro, Niterói, RJ, CEP:24020-140, Brazil.

PUC-Rio, Mechanical Engineering Departament, Rua Marquês de Sao Vicente, 225, Gávea, Rio de Janeiro, RJ, CEP:22451-900 , Brazil.

Université Paris-Est, Laboratoire Modelisation et Simulation Multi Echelle, MSME UMR 8208 CNRS, 5, Bd Descartes, 77454, Marne-La-Vallée, France

E-mail: ecataldo@vm.uff.br, rsampaio@puc-rio.br, christian.soize@univ-paris-est.fr

Abstract. The aim of this paper is to use Bayesian statistics to update a probability density function related to the tension parameter, which is one of the main parameters responsible for the changing of the fundamental frequency of a voice signal, generated by a mechanical/mathematical model for producing voiced sounds. We follow a parametric approach for stochastic modeling, which requires the adoption of random variables to represent the uncertain parameters present in the cited model. For each random variable, a probability density function is constructed using the Maximum Entropy Principle and the Monte Carlo method is used to generate voice signals as the output of the model. Then, a probability density function of the voice fundamental frequency is constructed. The random variables are fit to experimental data so that the probability density function of the fundamental frequency obtained by the model can be as near as possible of a probability density function obtained from experimental data. New values are obtained experimentally for the fundamental frequency and they are used to update the probability density function of the tension parameter, via Bayes's Theorem.

1. Introduction

One can say that the production of voiced sounds (vowels are particular cases of voiced sounds) starts with the contraction-expansion of the lungs causing an airflow, between the lungs and the mouth due to the difference of pressure, which will induce the auto-oscillation of the vocal folds (located in the larynx). After passing through the glottis and due to the movement of the vocal folds, the airflow is transformed into pulses of air which are generated (quasi)-periodically. The pressure signal created is so called the *glottal signal*, which will further be filtered and amplified by the vocal tract generating the sound we hear. The fundamental frequency of the voice signal, which is the frequency of the vocal folds oscillation, is the inverse of the period of the glottal signal.

Some authors have modeled the vocal folds dynamics, mainly in a deterministic way [8, 15, 7, ?]. One of these models is the well known model proposed by Ishizaka and Flanagan [4] and it will be used here because it has provided a simple and effective representation of the system to study the underlying dynamics of voice production. A diagram of the model can be seen in Fig 1.

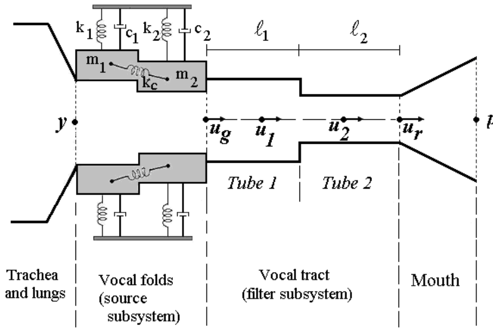

Figure 1. Two-mass model of the vocal folds.

The dynamics of the system is given by Eqs. (1) and (2) [1]:

$$\psi_1(\mathbf{w})\dot{u}_g + \psi_2(\mathbf{w})|u_g|u_g + \psi_3(\mathbf{w})u_g + \frac{1}{\tilde{c}_1}\int_0^t (u_g(\tau) - u_1(\tau))d\tau - y = 0 \tag{1}$$

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

where $\mathbf{w}(t) = (x_1(t), x_2(t), u_1(t), u_2(t), u_r(t))^t$, the functions x_1 and x_2 are the displacements of the masses, u_1 and u_2 describe the air volume flow through the (two) tubes that model the vocal tract and u_r is the air volume flow through the mouth. The subglottal pressure is denoted by y and u_g is the function that represent the glottal pulses signal. The function output radiated pressure p_r is given by $p_r(t) = u_r(t)r_r$, in which $r_r = \frac{128\rho v_c}{9\pi^3 y_2^2}$, ρ is the air density, v_c is the sound velocity, and y_2 is the radius of the second tube. The equations of the other quantities that appear in the equation, a detailed discussion of the model and its implementation can be found in [1].

As we can see, the process to generate a voiced sound is complex and its modeling involves and a lot of quantities which should be controlled. However, we are interested in the changing of the fundamental frequency. The main parameters responsible for these changing, as discussed in [4, 2, 1], are described in the following:

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 because vocal fold abduction and tension are the main factors used by a speaker to control phonation. In order to control the fundamental frequency of the vocal folds, parameters m_1, k_1, m_2, k_2, k_c 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.

These three parameters will be considered as uncertain and random variables will be associated to them. It means that for each realization of the three random variables a different voice signal is produced, characterizing that the voice production process generates a stochastic process.

The probability density functions associated to the random variables corresponding to the chosen uncertain parameters will be constructed by using the Maximum Entropy Principle (or better, the Jaynes's Maximum Entropy Principle).

The measure of uncertainty (entropy) used here was proposed by [11] and it is given by Eq.3:

$$S(p_X) = -\int_{-\infty}^{+\infty} p_X(x) ln\left(\, p_X(x)\,\right) dx\,. \tag{3}$$

The goal is to maximize the measure $S(p_X)$, subject to the constraints

$$\int_{-\infty}^{+\infty} p_X(x)dx = 1 \;\text{ and }\; \int_{-\infty}^{+\infty} p_X(x) g_i(x)dx = a_i\;,\; i = 1, \ldots, m \tag{4}$$

where a_i are usable information related to the functions g_i.

Then, according to the first part of the principle, we should use only probability distributions consistent with the constraints given. However, it may exist an infinity of probability distributions compatible with the constraints. The second part of the principle enable us to choose one among the many that satisfies the constraints, the (unique) probability distribution that maximizes the entropy.

This principle can be used for parametric probabilistic approach and also for nonparametric probabilistic approach [14].

2. Probabilistic model of the uncertain parameters

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 given 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 [12, 13]. Since no information is available concerning cross statistical moments between random variables A_{g0}, Y, Q, the random variable will be considered independent. The details related to the construction of the probability density functions related to these three random variables can be found in [1]. The expressions of the probability density functions will be described in the following.

The probability density function for the neutral glottal area is given by Eq.(5):

$$p_{A_{g0}}(a_{g0}) = \mathbf{1}_{]\mathbf{0},+\infty[} e^{-\lambda_0 - \lambda_1 a_{g0} - \lambda_2 (a_{g0})^2}, \tag{5}$$

where λ_0, λ_1 and λ_2 are the solution of the three equations defined by Eq.(8).

$$\int_{-\infty}^{+\infty} p_{A_{g0}}(a_{g0})\, da_{g0} = 1 \quad , \tag{6}$$

$$\int_{-\infty}^{+\infty} a_{g0}\, p_{A_{g0}}(a_{g0})\, da_{g0} = \underline{A}_{g0} \quad , \tag{7}$$

$$\int_{-\infty}^{+\infty} a_{g0}^2\, p_{A_{g0}}(a_{g0})\, da_{g0} = c, \tag{8}$$

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 $\delta_{A_{g0}^2} = E\{A_{g0}^2\}/\underline{A}_{g0}^2 - 1$ which proves that $c = \underline{A}_{g0}{}^2 \left(1 + \delta_{A_{g0}}^2\right)$.

The probability density function for the subglottal pressure is given by Eq.(9):

$$p_Y(y) = \mathbf{1}_{]0,+\infty[}(y) \frac{1}{\underline{Y}} \left(\frac{1}{\delta_Y^2}\right)^{\frac{1}{\delta_Y^2}} \times \frac{1}{\Gamma\left(1/\delta_Y^2\right)} \left(\frac{y}{\underline{Y}}\right)^{\frac{1}{\delta_Y^2}-1} exp\left(-\frac{y}{\delta_Y^2 \underline{Y}}\right)\,, \tag{9}$$

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

and the probability density function for the tension parameter is given by Eq.(10):

$$p_Q(q) = \mathbf{1}_{]0,+\infty[}(q) \frac{1}{\underline{Q}} \left(\frac{1}{\delta_Q^2}\right)^{\frac{1}{\delta_Q^2}} \times \frac{1}{\Gamma\left(1/\delta_Q^2\right)} \left(\frac{q}{\underline{Q}}\right)^{\frac{1}{\delta_Q^2}-1} exp\left(-\frac{q}{\delta_Q^2 \underline{Q}}\right) , \tag{10}$$

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. (10), it can be verified that Q is a second-order random variable and that $E\{1/Q^2\} < +\infty$.

3. Stochastic system

The equations for the deterministic case have been developed, the uncertain parameters have been identified and their probability density functions have been constructed. As explained above, the stochastic system is deduced from the deterministic one substituting a_{g0}, y, q by random variables A_{g0}, Y, Q. Consequently, 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) and (2) substituting a_{g0}, y, q by random variables A_{g0}, Y, Q. It will be calculated from the glottal signal, given by u_g, finding the inverse of its period.

In order to validate the development presented here, voice signals produced by one person have been analyzed and their statistics have been compared with simulations.

A voice signal corresponding to a sustained vowel /a/ has been recorded from one person and the function *fxrapt*, from MATLAB, was used to compute all of the time intervals corresponding to the opening and closure of the vocal folds were evaluated. For each time interval, a corresponding fundamental frequency was calculated as the inverse of the corresponding time interval. So, we can construct a corresponding probability density function that we will call *experimental*. Our goal is to solve an inverse problem in order to construct, from simulations, a probability density function similar to the experimental one. That is, we want to identify the mean values $\underline{A}_{g0}$, $\underline{Y}$, $\underline{Q}$, and also the coefficients of dispersion $\delta_{A_{g0}}$, δ_Y, δ_Q such that we can achieve the same *experimental* mean value of the fundamental frequency $m_{F_0} = 120.7694$ Hz and also the same *experimental* coefficient of dispersion of the fundamental frequency $\delta_{F_0} = \dfrac{m_{F_0}}{\delta_{F_0}} = 0.0173$. We want also to compare the shapes of the distributions: experimental and constructed. This inverse problem has not necessarily an unique solution. So, we will present an example, considering two different possible solutions of the inverse problem. The strategy used will be described in the following:

Step 1: Values of a_{g_0}, y, and q are chosen, in the corresponding deterministic problem, such that an output radiated pressure signal with fundamental frequency $f_0 = 120.7694$ Hz is obtained.

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 $\underline{Q}$ in the corresponding stochastic problem.

Step 3: With the mean values described in Step 2, values of $\delta_{A_{g0}}$, δ_Y, and δ_Q are chosen such that the value of $\delta_{F_0} = \dfrac{m_{F_0}}{\sigma_{F_0}} = 0.0173$. A Monte Carlo method is used.

Clearly, in order to identify the parameters as described many tests were made. If the number of cases is large, a strategy to solve this inverse problem, for example, creating an adequate cost function, might be more effective.

The values obtained for each *step* were:

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

Step 2: The values $\underline{A}_{g0} = 5 \times 10^{-2}\,\text{m}^2$, $\underline{Y} = 750\,\text{Pa}$, and $\underline{Q} = 0.66$ are used in the corresponding stochastic problem.

Step 3: With the mean values described in Step 2, the mean value of the fundamental frequency obtained, considering 700 realizations and using the Monte Carlo method, was $m_{F_0} = 120.9525\,\text{Hz}$. With the values of the coefficients of dispersion $\delta_{A_{g0}} = 0.03$, $\delta_Y = 0.02$, $\delta_Q = 0.02$, the value obtained for the coefficient of dispersion of the fundamental frequency was $\delta_{F_0} = 0.0171$.

Figure 2 shows the probability density function constructed from experimental signals (top) and the probability density function constructed from simulations (bottom).

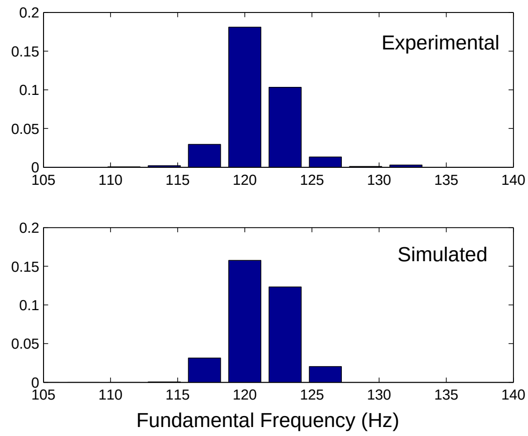

Figure 2. Histogram of the fundamental frequency distribution: experimental (top) and simulated (bottom).

Then, it shows that the mean values and the dispersion values used for the random variables Q, Ag_0 and Y were coherent. From the histograms, it can be constructed the probability density functions. It will be used the function *ksdensity* from MATLAB and the plots are shown in Fig 3.

In order to better fit the data, it will be used bayesian statistics to update the probability density function of the fundamental frequency.

4. Bayesian statistics applied to the problem

It is used the parametric probabilistic approach to take into account the uncertainties related to the parameters of the model. The probabilistic model is constructed in two steps: first, it is constructed the prior probabilistic model using the Maximum Entropy Principle in the context of Information Theory. Second, it is constructed a posterior probabilistic model using experimental data and the Bayes method.

The idea is to update the probability distribution of the random variable Q. Only one parameter was chosen because the results can be better controlled and analyzed. This parameter, Q, is the most important by the changing of the fundamental frequency and this was the reason for what it was chosen. Moreover, it is a parameter that describes the tension of the vocal folds and it is very difficult to obtain experimental data for it.

Let F_0 be the fundamental frequency of the voice signal, $h(q)$ the prior density probability function for Q and $f(f_0|q)$ the likelihood function. Then, using the Bayes's theorem for probability density functions one can write the Eq. 11.

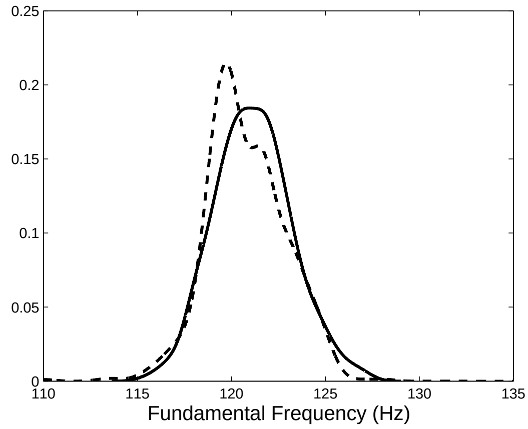

Figure 3. Probability density functions: experimental (dashed line) and simulated (continuous line).

$$h(q|f_0) = \frac{f(f_0|q)h(q)}{\displaystyle\int_Q f(f_0|q)h(q)dq} . \quad (11)$$

The the prior probability density function $h(Q)$ is given by Eq. 12.

$$h(q) = \mathbf{1}_{]0,+\infty[}(q)\frac{1}{\underline{Q}}\left(\frac{1}{\delta_Q^2}\right)^{\frac{1}{\delta_Q^2}} \frac{1}{\Gamma\left(1/\delta_Q^2\right)} \left(\frac{q}{\underline{Q}}\right)^{\frac{1}{\delta_Q^2}-1} exp\left(-\frac{q}{\delta_Q^2\underline{Q}}\right) . \quad (12)$$

Figure 4 shows the plot of the prior probability density function $h(Q)$, considering the values $\underline{Q} = 0.66$ and $\delta_Q = 0.02$.

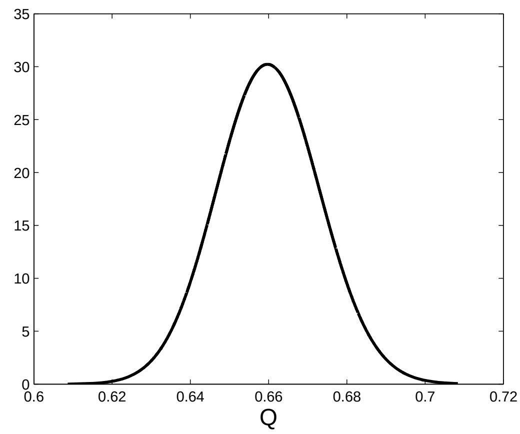

Figure 4. Prior probability density function for Q.

The likelihood function $f(f_0|\theta)$ is the one obtained numerically, by simulation, considering the values discussed above.

Then, a new experimental value was obtained for the fundamental frequency (it was 122 Hz). With this value, the prior density probability function for Q was updated, using Eq.(11) and its plot is shown in Fig.(5).

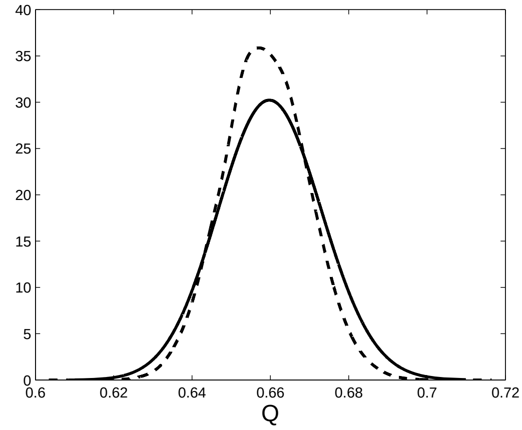

Figure 5. Prior probability density function ($h(q)$) (full line) and Posterior probability density function ($h(q|f_0)$ (dashed line).

Another value is get for the experimental fundamental frequency (at this time, 122.5 Hz) and once more the prior probability density function for Q is updated. The plots for the approximations of the three density probability functions for Q are showed in the Fig. 6.

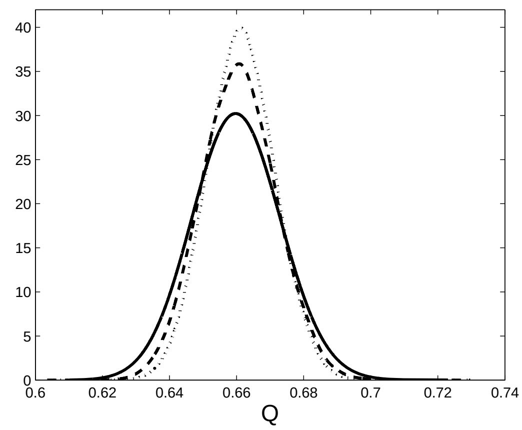

Figure 6. Third iteration for the posterior probability density function (line with points, the higher).

With the new probability density function found for Q, new values were simulated for the fundamental frequency and another probability density function for the fundamental frequency was generated. Its plot is added to the Fig. 3 and showed in the Fig. 7.

5. Conclusions

Using Bayesian statistics, the probability density function of the random variable Q related to an important parameter which takes part in a mathematical model for producing voice, was

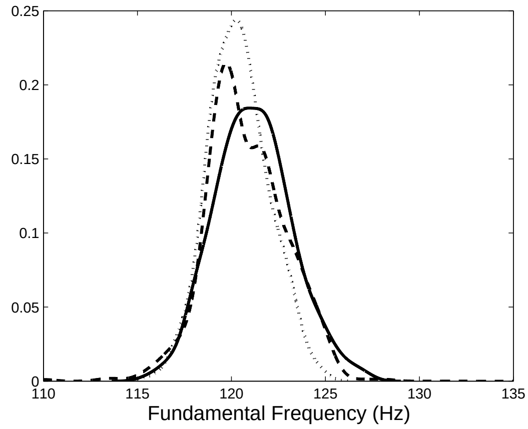

Figure 7. Probability density functions: experimental (dashed line), first simulation (continuous line) and second simulation (line with points, the higher).

updated after obtaining new experimental data. It should be observed that the first prior probability density function for Q was obtained using the Maximum Entropy Principle, and there is difficulty to obtain real values for this parameter, because it is related to a biological quantity. Using Bayes's Theorem it was possible to update the probability density function without getting values directly for this parameter, but from other quantity (the fundamental frequency) which can be easily observed. From the posterior probability density function obtained for Q, it was possible to simulate voice signals and to construct a posterior probability density function for the fundamental frequency. This new p.d.f. was compared with the probability density function of the fundamental frequency obtained experimentally and, as observed, the both p.d.f.'s were nearer, better fitted.

Acknowledgments
This work was supported by Fundação de Amparo à Pesquisa no Rio de Janeiro (FAPERJ - programa Jovens Cientista do Nosso Estado), by CAPES (CAPES/COFECUB project N. 672/10) and by the Brazilian Agency Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

References
[1] Cataldo E, Soize C, Sampaio R and Desceliers C 2009 probabilistic modeling of a nonlinear dynamical system used for producing voice *Computational Mechanics* **43** 265–275.
[2] Cataldo E, Sampaio R, Lucero J and Soize C 2008 modeling random uncertainties in voice production using a parametric approach. *Mechanics Research Communication*, **35**, 429–490.
[3] Herzel H 1993 bifurcations and chaos in voice signals *Applied Mechanics Review*, **7**, 399-411.
[4] Ishizaka K and Flanagan J LK. Ishizaka 1972 synthesis of voiced sounds from a two-mass model of the vocal cords *Bell Syst. Tech. J.*, **51**, 1233-1268.
[5] Jaynes E 1957 information theory and statistical mechanics *Phys. Rev.*, **106(4)**, 620-630.
[6] Jaynes E 1957 Information theory and statistical mechanics *Phys. Rev.*, **108**, 171-190.
[7] Koizumi T, Taniguchi S, Hiromitsu S 1976 two-mass models of the vocal cords for natural sounding voice synthesis, *J. Acoust. Soc. Am.*, **82**, 1179-1183.
[8] Lous N J C, Hofmans G C J, Veldhuis R N J and Hirschberg A 1998 symmetrical two-mass vocal-fold model coupled to vocal tract and trachea with application to prosthesis design *Acta Acustica*, **84 (6)**, 1135-1150.
[9] Mauprivez J, Sampaio R and Cataldo E 2009 parameters setting through optimal experimental design for a two-mass vocal folds model *11th Pan-American Congress of Applied Mechanics - PACAM XI*

[10] Paulino C D, Turkman M A A and Murteira B 2003 estatística bayesiana *Fundação Clouste Gulbenkian Lisboa*,
[11] Shannon C E 1948 a mathematical theory of communication *Bell System Tech*, **28**, 379-423 and 623-659.
[12] Soize C 2000 a nonparametric model of random uncertainties for reduced matrix models in structural dynamics *Probabilistic Engineering Mechanics*, **15**, 277-294.
[13] Soize C 2001 maximum entropy approach for modeling random uncertainties in trasient elastodynamics *Journal of the Acoustical Society of America*, **109(5)**, 1979-1996.
[14] Soize C 2010 generalized probabilistic approach of uncertainties in computational dynamics using random matrices and polynomial chaos decompositions. *International Journal for Numerical Methods in Engineering*, **I.2,IV.1,IV.2**
[15] Zhang Y, Jiang J and Rahn III D A 2005 studying vocal fold vibrations in Parkinson's disease with a nonlinear model *Chaos* **15**, 1-10.