Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Separable Least Squares Identification of Long Memory Block Structured Models: Application to Lung Tissue Viscoelasticity

David T. Westwick
Department of Electrical and Computer Engineering
Schulich School of Engineering, Calgary, AB, Canada

Bela Suki
Department of Biomedical Engineering
Boston University, Boston, MA, USA

***Abstract*— A separable least squares algorithm is developed for the identification of a Wiener model whose dynamic element is a constant phase model that has been modified to include a purely viscous term. The separation of variables reduces the dimensionality of the search space from 5 to 2, greatly simplifying the optimization procedure used to estimate the parameters. The algorithm is tested on experimental stress/strain data from a strip of lung parenchyma.**

***Index Terms*— Nonlinear System, Power Law, Optimization, Tissue Strips, Stress Relaxation, Constant Phase Model.**

I. Introduction

Stress relaxation behavior of a tissue is defined to be the dynamic relationship between an input strain, and the resulting stress. Since strain and stress are proportional to displacement and force, respectively, the stress relaxation can be thought of as a normalization of the dynamic stiffness.

The stress relaxation of lung tissue has been observed to have a very long memory. Indeed, it has been suggested [1], [2] that the decay of the stress should resemble a fractional integrator, or power-law relationship. Such power-law relationships have been observed experimentally, both for whole lungs [2], and for strips of lung tissue [3]. Since the responses of power-law systems eventually decay more slowly than any exponential, they are often called "long memory systems".

Given the complexity of the tissue, it is unlikely that any mathematical model derived from first principles will be of any practical use. System identification techniques [4], however, may be used to obtain such a model from experimental data. Indeed, in two separate studies [3], [5], the dynamics of stress relaxation behavior in lung tissue strips have been identified from experimental data. These studies both used simple block structured models, interconnections of dynamic linear systems and zero-memory nonlinearities, to capture the dynamic nonlinear behavior of the tissue strip system. They both considered the two simplest block structures: the Wiener cascade, a linear dynamic element followed by a zero-memory nonlinearity, and the Hammerstein cascade, a zero-memory nonlinearity followed by a dynamic linear system [6].

In Maksym *et al.* [5], the linear dynamic elements were represented using finite impulse response (FIR) filters, while polynomials were used for the static nonlinearities. The Hammerstein structure was found to provide more accurate predictions of the measured stress than the Wiener structure.

Yuan *et al.* [3] also considered both Wiener and Hammerstein models, and also used polynomials to represent the zero-memory nonlinearities. However, they considered a variety of representations for the linear dynamic element: nonparametric FIR filters, nonparametric frequency domain models, and parametric frequency-domain models. They found that the best predictions were obtained from the models which used a parametric representation of the linear element. Furthermore, the best overall predictions were obtained from the parametric Wiener structure.

The parametric models considered by Yuan *et al.* included the power-law dynamics explicitly [3]. Unlike the FIR filters, where the number of "parameters" was proportional to the memory length, the parametric models depended on only 3 parameters. Hence, when a 3'rd order polynomial was used as the nonlinearity, the parametric tissue model included 5 parameters: the 3 that described the linear power-law dynamics, and two variable polynomial coefficients.

A global nonlinear optimization [7] was used to find the optimal parameters in the 5 dimensional search space. This algorithm split the search space into segments, and a gradient-based optimization was initiated at the center of each segment. The results of the initial optimizations were clustered, and the center of the best cluster was used as the starting point for the final optimization. This approach increases the likelihood that the final solution is the global optimum.

In this paper, we will show how a new approach to block structured system identification, separable least squares optimization [8], can be used to reduce the dimension of the nonlinear search space from 5, to 2. As a result, the global optimization scheme need only consider a 2 dimensional space, greatly reducing the number of starting points, and hence iterative local optimizations, that need to be performed. The shape of the resulting two-dimensional error surface will be examined, and the basin of attraction surrounding the globally optimal solution will be characterized.

II. Analytical Methods

In this paper, upper and lower case letters will be used to denote matrices and vectors, respectively. Elements of vectors and matrices will be given the same symbol, without the bolding, with the indices in parentheses. Thus $\theta(j)$ will represent the j'th element of the vector $\boldsymbol{\theta}$. The superscripts

T and $\ddagger$ will be used to denote transposition, and the Moore-Penrose pseudo-inverse respectively,

$$\mathbf{A}^{\ddagger} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$$

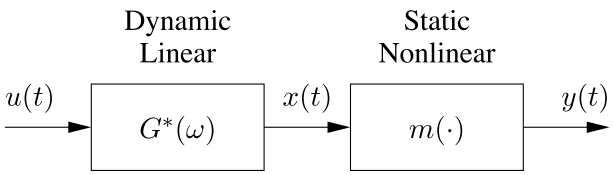

Fig. 1. Block diagram of a Wiener cascade model of stress relaxation. The input $u(t)$ is the strain, $x(t)$ is an unmeasurable internal signal, and $y(t)$ represents the output stress. Note that the dynamic linear element is described in the frequency domain, while the static nonlinearity is described in the time domain.

A. Parametric Wiener Models

The Wiener model comprises a linear dynamic element followed by a zero-memory, or static, nonlinearity, as shown in Fig. 1. [6]. Several authors [3], [9] have represented the complex modulus of lung tissue dynamics, both of whole lungs and of tissue strips, in the frequency domain, using the frequency response,

$$G^*(\omega) = H\omega^{\beta} + j\left(G\omega^{\beta} + R\omega\right) \tag{1}$$

where the parameters G, H and β are related by

$$\beta = 1 - \frac{2}{\pi}\tan^{-1}\left(\frac{H}{G}\right) \tag{2}$$

Given the relationship between G, H and β, there are only two independent parameters. For the purposes of estimation, it will be more convenient to work with the model

$$G^*(\omega) = Qe^{j\frac{\pi}{2}\beta}\omega^{\beta} + jR\omega \tag{3}$$

where $Q^2 = G^2 + H^2$.

The static nonlinearity will be modeled using a third-degree polynomial, as used in [3]:

$$y(t) = m(x(t)) = \sum_{\ell=1}^{3} c_{\ell}x(t)^{\ell} \tag{4}$$

The Wiener cascade, consisting of the frequency domain description of the linear dynamics (3) and the time-domain description of the memoryless nonlinearity (4), can be completely specified by the 3 parameters that describe the linear dynamics, Q, R and β, and the 3 polynomial coefficients, $c_1 \ldots c_3$. Note, however, that one of these parameters is redundant [6], [3], since any gain applied to the linear dynamics can be canceled by rescaling the polynomial coefficients. In Yuan *et al*. [3], this redundancy was was eliminated by fixing the first order polynomial coefficient, $c_1 = 1$. In the sequel, it will be advantageous to fix the magnitude of the dynamic linear element, rather than the of the nonlinearity. Thus, we will set $Q = 1$, so that the linear dynamics are now described by only two parameters,

$$G^*(\omega) = e^{j\frac{\pi}{2}(1-\beta)}\omega^{\beta} + jR\omega \tag{5}$$

but, as a result, all of the polynomial coefficients will be free parameters. Thus, the Wiener cascade can be represented by the parameter vector,

$$\boldsymbol{\theta} = \left[\begin{array}{cc|ccc} \beta & R & c_1 & c_2 & c_3 \end{array}\right]^T \tag{6}$$

B. Separable Least Squares Identification

The parameter vector $\boldsymbol{\theta}$, defined in (6), fully defines the Wiener model. The output of the model defined by the parameter vector $\boldsymbol{\theta}$ will be written as $\hat{y}(t, \boldsymbol{\theta})$. The objective in a minimum mean-squared system identification is to find the parameter vector that minimizes the cost function

$$V_N(\boldsymbol{\theta}) = \frac{1}{2N}\sum_{t=1}^{N}\left(y(t) - \hat{y}(t, \boldsymbol{\theta})\right)^2$$

In many cases, the output of a model depends linearly on some of its parameters, but nonlinearly on the rest. In these cases, it is possible to formulate the system identification as a separable least squares (SLS) optimization [10], [11] problem. SLS methods have been used in the identification of both Hammerstein [8], [12] and Wiener [13], [14] systems, using conventional FIR and IIR digital filter representations of the linear elements. A similar approach may be applied to the long memory systems considered here.

Consider the Wiener structure, shown in Fig. 1, where the linear element is the modified constant phase model (1) and the nonlinearity is represented by a third degree polynomial (4). Divide the parameter vector, $\boldsymbol{\theta}$, defined in (6), into two segments: $\boldsymbol{\theta_l}$, which contains the parameters that enter linearly into the output, and $\boldsymbol{\theta_n}$, containing the parameters that enter nonlinearly. Thus,

$$\begin{aligned} \boldsymbol{\theta} &= \left[\begin{array}{cc} \boldsymbol{\theta_n}^T & \boldsymbol{\theta_l}^T \end{array}\right]^T \\ \boldsymbol{\theta_n} &= \left[\begin{array}{cc} \beta & R \end{array}\right]^T \\ \boldsymbol{\theta_l} &= \left[\begin{array}{ccc} c_1 & c_2 & c_3 \end{array}\right]^T \end{aligned}$$

The output of the linear element depends only on $\boldsymbol{\theta_n}$, the nonlinear parameters,

$$\hat{x}(t, \boldsymbol{\theta_n}) = G(q, \boldsymbol{\theta_n})u(t) \tag{7}$$

Let $\mathbf{X}(\boldsymbol{\theta_n})$ be a $N \times 3$ matrix whose columns contain the linear system output raised to the first, second and third powers, respectively. Then, the output of the Wiener system is given by:

$$\hat{\mathbf{y}}(\boldsymbol{\theta}) = \mathbf{X}(\boldsymbol{\theta_n})\boldsymbol{\theta_l} \tag{8}$$

The mean squared error is

$$V_N(\boldsymbol{\theta}) = \frac{1}{2N} \parallel \mathbf{y} - \mathbf{X}(\boldsymbol{\theta_n})\boldsymbol{\theta_l} \parallel_2^2 \tag{9}$$

For any given value of $\boldsymbol{\theta_n}$, the model output is a linear function of the polynomial coefficients. Therefore, for any given $\boldsymbol{\theta_n}$, the optimal polynomial coefficients are obtained by solving the linear regression,

$$\hat{\boldsymbol{\theta}}_{\boldsymbol{l}} = \mathbf{X}(\boldsymbol{\theta_n})^{\ddagger}\mathbf{y} \tag{10}$$

Since the (locally optimal) polynomial coefficients are functions of $\boldsymbol{\theta_n}$, the model output and cost function can both be viewed as functions of $\boldsymbol{\theta_n}$ alone. Thus, the 5 dimensional search space has been transformed into a 2 dimensional search space.

Iterative nonlinear minimization algorithms, such as the Gauss-Newton and Levenberg-Marquardt algorithms [15], require computing the Jacobian, a matrix whose columns contain the partial derivatives of model output with respect to each of the parameters. To apply any of these optimization algorithms to the separated problem, we must compute the Jacobian of the separated problem,

$$\mathbf{J_s}(t,k) = \left.\frac{\partial \hat{y}(t,\boldsymbol{\theta}_n)}{\theta_n(k)}\right|_{\boldsymbol{\theta_l}=\hat{\boldsymbol{\theta}}_{\boldsymbol{l}}(\boldsymbol{\theta_n})}$$

The columns of $\mathbf{J_s}$ contain the partial derivatives of the model output with respect to the nonlinear parameters, under the assumption that the linear parameters continually take on their locally optimal values. This can be computed from the Jacobian of the unseparated problem using the variable projection method [11]. Let

$$\begin{aligned} J_l(t,k) &= \frac{\partial \hat{y}(t,\boldsymbol{\theta})}{\theta_{l,k}} \\ J_n(t,k) &= \frac{\partial \hat{y}(t,\boldsymbol{\theta})}{\theta_{n,k}} \end{aligned}$$

be the partial derivatives of the model output with respect to the linear and nonlinear parameters, respectively, under the assumption that all other parameters remain fixed. Then, the Jacobian of the separated problem is given by [8], [11]

$$\mathbf{J_s} = \mathbf{J_n} - \mathbf{J_l}\mathbf{J_l}^{\ddagger}\mathbf{J_n}$$

which is the projection of $\mathbf{J_n}$ orthogonal to the columns of $\mathbf{J_l}$. Once $\mathbf{J_s}$ has been computed, the optimization can be performed using the standard Gauss-Newton or Levenberg-Marquardt techniques.

C. Separation of Parametric Wiener Models

The output of the Wiener cascade is given by (8). Thus, the linear Jacobian is

$$\mathbf{J_l} = \mathbf{X}(\boldsymbol{\theta_n}) \tag{11}$$

From [16], the k'th column of the nonlinear Jacobian is given by

$$J_n(t,k) = \frac{\partial \hat{x}(t,\boldsymbol{\theta_n})}{\partial \theta_n(k)} m'(\hat{x}(t,\boldsymbol{\theta_n})) \tag{12}$$

where m' is the derivative of the static nonlinearity with respect to its input,

$$m'(x) = \sum_{\ell=1}^{3} \ell c_\ell x^{\ell-1}$$

Computing the partial derivatives of $\hat{x}(t,\boldsymbol{\theta_n})$ with respect to β and R is most easily done in the frequency domain:

$$\begin{aligned} \frac{\partial}{\partial R}\hat{X}(\omega,\boldsymbol{\theta_n}) &= j\omega U(\omega) && (13) \\ \frac{\partial}{\partial \beta}\hat{X}(\omega,\boldsymbol{\theta_n}) &= \left(\log(\omega) + \frac{j\pi}{2}\right)\hat{X}(\omega,\boldsymbol{\theta_n}) && (14) \end{aligned}$$

These can then be inverse Fourier transformed and inserted into (12) to yield the columns of $\mathbf{J_n}$.

III. Experimental Methods

The experiment was similar to that reported in [3], [17], but used an updated apparatus [18]. Briefly, a 1.9x1.9x2 mm strip of Guinea Pig parehchyma was suspended in a glass tissue bath, between a force transducer (model 403A, Aurora Scientific, Ontario, Canada) and a position controlled servo arm (model 300B, Aurora Scientific). The position servo imposed a constant stretch equal to 40% of the resting length of the tissue strip. A non-sum non-difference multi-sine input [19], with a peak to peak amplitude equal to 20% of the sample's resting length, was then superimposed on the constant baseline stretch. The input position and output force signals were then low-pass filtered at 15 Hz (901P Filter Bank, Frequency Devices, Haverhill, MA), and sampled at 60 Hz using a data acquisition card (DAQCard6062E, National Instruments). Each trial lasted 68 seconds, corresponding to 4096 input/output data points. The displacement and force signals were normalized by the reference length and cross-sectional area of the strip, to produce strain and stress, respectively.

IV. Results

After separation of variables, the Wiener model has only two free parameters. As a result, the sum of squared errors cost function may be plotted in three dimensions as a surface. Figure 2 shows the cost function plotted versus β and R, setting $Q = 1$ to eliminate the redundant gain in the model, and allowing the polynomial coefficients to take on their optimal values corresponding to each set of (β, R) values. For $0 < \beta < 1$, and $-1 < R < 1$, the error surface appears to have a single minimum near $\beta = 0.05$ and $R = 0.01$. Although the minimum is not evident in Fig. 2, it can be seen in Fig. 3, which shows the minimum attainable cost, as a function of the exponent β.

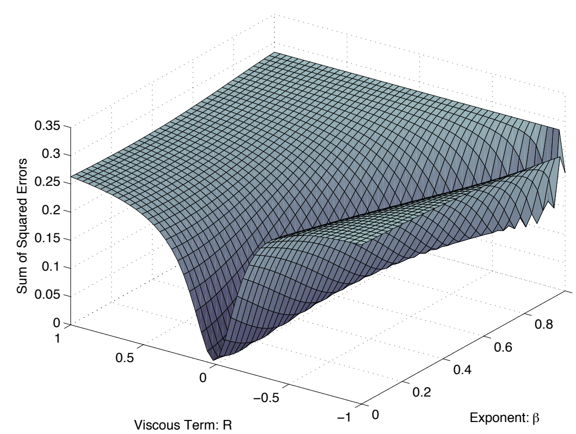

Fig. 2. Sum of squared errors for a long-memory Wiener model with a degree 3 nonlinearity. The SSE is plotted as a function of β and R, assuming that $Q = 1$, and the optimal polynomial coefficients are used for each β and R. Thus, the value on the y-axis indicates the relative magnitudes of the viscous (R) and constant phase (Q) terms.

Indeed, starting a Levenberg-Marquardt optimization at any point in this range produces a final solution with $\beta =$

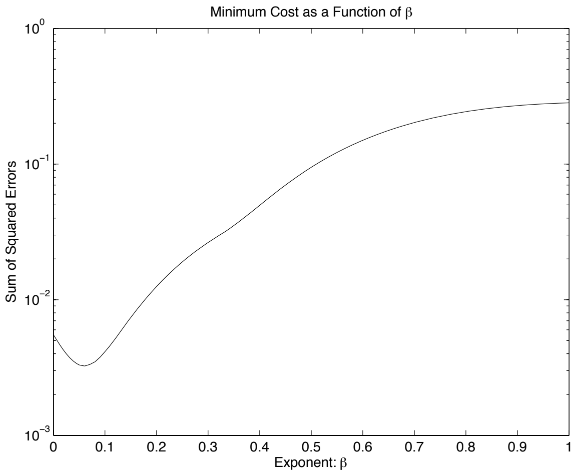

Fig. 3. Sum of squared errors as a function of the exponent β, assuming that the optimal value of R has been chosen. This corresponds to the bottom of the valley evident in Fig. 2.

0.0575 and $R = 0.0028$. Rescaling these values so that the linear polynomial coefficient is 1, and transforming the constant-phase portion of the model yields $G = 0.0174$ and $H = 0.1919$. In this case, the basin of attraction for the global optimum was very large, and encompassed all physiologically relevant starting points.

V. Discussion

The primary contribution of this paper is the development of a separable least squares optimization algorithm for fitting Wiener models whose linear dynamics can be described by the sum of a constant phase model and a purely viscous term. Although the memoryless nonlinearity must be analyzed in the time-domain, the linear dynamics, which contain a differentiator due to the viscous term and a fractional differentiator from the constant-phase component, are most easily handled in the frequency domain.

Like previous algorithms, this approach requires the use of an iterative nonlinear optimization. However in this case, the optimization is performed on a two-dimensional parameter space, regardless of the degree of the polynomial nonlinearity. Furthermore, in the data set analyzed in this paper, and in other datasets obtained from the same preparation, the error surface appeared to have a well characterized global minimum surrounded by a large basin of attraction, thus eliminating the need for a sophisticated, time-consuming global optimization, as used in Yuan *et al*. [3]. Even if a global optimization should become necessary, it could be performed over the two dimensional parameter space corresponding to the separated model, rather than the full parameter space. Finally, as with previous parametric optimization based solutions, the separable least squares algorithm proposed in this paper is not limited to Gaussian white noise test inputs, but may be used with any sufficiently rich input, such as the non-sum non-difference multi-sine used in the experiment.

Acknowledgment

Partially supported by NSERC (Canada) and NIH (USA). The authors would like to thank Mr. Lauren Black for the acquisition of the experimental data.

References

[1] J. Bates, G. Maksym, D. Navajas, and B. Suki, “Lung tissue reology and 1/f noise,” *Annals of Biomedical Engineering*, vol. 22, pp. 674–681, 1994.

[2] B. Suki, A. Barabási, and K. Lutchen, “Lung tissue viscoelasticity: a mathematical framework and its molecular basis,” *Journal of Applied Physiology*, vol. 76, no. 6, pp. 2749–2759, 1994.

[3] H. Yuan, D. Westwick, E. Ingenito, K. R. Lutchen, and B. Suki, “Parametric and nonparametric nonlinear system identification of lung tissue strip mechanics,” *Annals of Biomedical Engineering*, vol. 27, no. 4, pp. 548–562, 1999.

[4] D. Westwick and R. Kearney, “Nonparametric identification of nonlinear biomedical systems, part i: Theory,” *Critical Reviews in Biomedical Engineering*, vol. 26, no. 3, pp. 153–226, 1998.

[5] G. Maksym, R. Kearney, and J. Bates, “Nonparametric block-structured modeling of lung tissue strip mechanics,” *Annals of Biomedical Engineering*, vol. 26, no. 2, pp. 242–252, 1998.

[6] I. Hunter and M. Korenberg, “The identification of nonlinear biological systems: Wiener and Hammerstein cascade models,” *Biological Cybernetics*, vol. 55, pp. 135–144, 1986.

[7] T. Csendes and D. Ratz, “Subdivision direction selection in interval methods for global optimization,” *SIAM Journal of Numerical Analysis*, vol. 34, no. 3, pp. 922–938, 1997.

[8] D. Westwick and R. Kearney, “Separable least squares identification of nonlinear Hammerstein models: Application to stretch reflex dynamics,” *Annals of Biomedical Engineering*, vol. 29, no. 8, pp. 707–718, August 2001.

[9] B. Suki, Q. Zhang, and K. Lutchen, “Relationship between frequency and amplitude dependece in the lung: a nonlinear block-structured modeling approach,” *Journal of Applied Physiology*, vol. 79, no. 2, pp. 660–671, 1995.

[10] G. Golub and V. Pereyra, “The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate,” *SIAM Journal of Numerical Analysis*, vol. 10, no. 2, pp. 413–432, 1973.

[11] J. Sjöberg and M. Viberg, “Separable non-linear least squares minimization – possible improvements for neural net fitting,” in *IEEE Workshop on Neural Networks for Signal Processing*, vol. 7, 1997, pp. 345–354.

[12] E. Dempsey and D. Westwick, “Identification of Hammerstein models with cubic spline nonlinearities,” *IEEE Transactions on Biomedical Engineering*, vol. 51, pp. 237–245, 2004.

[13] M. Hughes and D. Westwick, “Identification of iir wiener systems with spline nonlinearities that have variable knots,” *IEEE Transactions on Automatic Control*, vol. 50, no. 10, pp. 1617–1622, 2005.

[14] J. Bruls, C. Chou, B. Haverkamp, and M. Verhaegen, “Linear and nonlinear system identification using separable least-squares,” *European Journal of Control*, vol. 5, no. 1, pp. 116–128, 1999.

[15] J. Nocedal and S. Wright, *Numerical Optimization*. New York: Springer-Verlag, 1999.

[16] D. Westwick and R. Kearney, *Identification of Nonlinear Physiological Systems*, ser. IEEE Press Series in Biomedical Engineering. Piscataway, NJ: IEEE Press / Wiley, 2003.

[17] H. Yuan, E. Ingenito, and B. Suki, “Dynamic properties of lung parenchyma: Mechanical contributions of fiber network and interstitial cells,” *Journal of Applied Physiology*, vol. 83, no. 5, pp. 1420–1431, 1997.

[18] L. Black, K. Brewer, S. Morris, B. Schreiber, P. Toselli, M. Nugent, B. Suki, and P. Stone, “Effects of elastase on the mechanical and failure properties of engineered elastin-rich matrices,” *Journal of Applied Physiology*, vol. 98, pp. 1434–1441, 2005.

[19] K. Lutchen, K. Yang, D. Kaczka, and B. Suki, “Optimal ventilation waveforms for estimating low-frequency respiratory impedance,” *Journal of Applied Physiology*, vol. 75, pp. 478–488, 1993.