Fast System Identification Algorithm for Non-uniformly Sampled Noisy Biomedical Signal

Koon-Pong Wong [1], Dagan Feng [2], and Wan-Chi Siu [1]

[1] Department of Electronic Engineering
The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

[2] Basser Department of Computer Science
The University of Sydney, N.S.W. 2006, Australia

ABSTRACT — *The recently developed generalized linear least squares (GLLS) algorithm has been found very useful in non-uniformly sampled biomedical signal processing and parameter estimation. In this paper, the algorithm is used for the identification of a compartmental model with a pair of repeated eigenvalues based on the non-uniformly sampled noisy data. A case study is presented, which demonstrates that the algorithm is able to select the most suitable model for the system from the non-uniformly sampled noisy signals.*

I. INTRODUCTION

System identification is generally referred to as the determination of a mathematical model for a system or a process by observing the input-output relationships. In biomedical system identification, the key issue is to estimate the physiological parameters of a model based on the sampled data from a dynamic process. This model can usually be described by a set of differential equations. If the data can be sampled uniformly as in most of the engineering systems, the *indirect method* can be utilized and the signal processing is relatively straightforward. The *indirect method* proceeds by first estimating the parameters for a discrete-time model that best fits the data at these uniformly sampled times. The discrete-time model parameters are subsequently converted into the parameters of an equivalent continuous-time model through Impulse-Invariance method, which guarantees that the continuous-time and discrete-time models have identical output values at these uniformly sampled instances. The advantage of using the *indirect method* is that a number of effective parameters estimation algorithms are available, e.g. Least Squares (LS), Instrumental Variable (IV), Multistage Least Squares, and Generalized Least Squares (GLS) [6]. Among these methods, GLS has the highest accuracy. GLS iteratively upgrades the estimates that are initially obtained from the Least Squares and the final estimates are unbiased.

However, in contrary to engineering systems, the dynamic data in most biomedical systems are usually sampled non-uniformly [7]. In these cases, the well-developed *indirect method* algorithms cannot be used. The *direct method*, on the other hand, estimates the continuous-time model parameters by fitting these non-uniformly sampled data directly. The classic *nonlinear least squares (NLS)* is widely applied and can provide parameter estimates of optimum statistical accuracy [10]. Nevertheless, good initial parameter values are required and the computational complexity of this algorithm is very high. If the initial parameter values are not close enough to the parameter true values, NLS will converge very slowly or even not converge at all. Other algorithms such as System Reference Adaptive Model, Maximum Likelihood (ML), and Prediction Error [2] are all very time-consuming. As a result, they are impractical for high resolution image-wide parameter estimation.

Recently, Feng *et al* [5] proposed an *Generalized Linear Least Squares (GLLS)* algorithm for parameter estimation of non-uniformly sampled biomedical systems. This algorithm can provide unbiased parameter estimation with very little computing time and without the need of providing the initial parameter values. As it is statistically reliable and computationally efficient, it has been found to be very useful for biomedical system identification and image-wide parameter estimation [4,5]. However, this algorithm cannot deal with the signals and systems containing repeated eigenvalues which often occur in biological systems [3]. Wong *et al* [9] extended it so that it can be used for identification of system containing repeated eigenvalues as well. In this paper, the GLLS algorithm is used for the identification of a compartment model and the case study demonstrates that the algorithm is able to select the suitable model for the system based on the non-uniformly sampled data.

II. THEORY

The general Single-Input-Single-Output (SISO) linear continuous dynamic system can be described by the following n-th order differential equation:

0-7803-3679-8/96/$5.00 © 1996 IEEE

Authorized licensed use limited to: Hong Kong Polytechnic University. Downloaded on June 25,2010 at 04:06:09 UTC from IEEE Xplore. Restrictions apply.

$$y^{(n)}(t) + a_1 y^{(n-1)}(t) + \ldots + a_n y(t) = b_1 u^{(n-1)}(t) + b_2 u^{(n-2)}(t) + \ldots + b_n u(t) \quad (1)$$

where u(t) and y(t) are the input and output of the system respectively, $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ are the system transfer function parameters. The Laplace transform of equation (1) is given by

$$s^n Y(s) - s^{n-1}y(0) - \ldots - y^{(n-1)}(0) + a_1[s^{n-1}Y(s) - s^{n-2}y(0) - \ldots - y^{(n-2)}(0)] + \ldots + a_n Y(s) = b_1[s^{n-1}U(s) - s^{n-2}u(0) - \ldots - u^{(n-2)}(0)] + \ldots + b_n U(s) \quad (2)$$

where $u(0), \ldots, u^{(n-2)}(0)$ and $y(0), \ldots, y^{(n-1)}(0)$ are the initial conditions for the input and output functions. Equation (2) can be rewritten as

$$(s^n + a_1 s^{n-1} + \ldots + a_n)Y(s) = (b_1 s^{n-1} + b_2 s^{n-2} + \ldots + b_n)U(s) + v_1 s^{n-1} + v_2 s^{n-2} + \ldots + v_n \quad (3)$$

where $v_1, v_2, \ldots, v_n$ are the linear combinations of the input and output initial conditions, with $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ as their coefficients. v_i (i = 0, 1, 2, ... , n) can be written in matrix form as below:

$$\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ a_1 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ a_{n-1} & a_{n-2} & \ldots & 1 \end{bmatrix} \begin{bmatrix} y(0) \\ y^{(1)}(0) \\ \vdots \\ y^{(n-1)}(0) \end{bmatrix} - \begin{bmatrix} 0 & 0 & \ldots & 0 \\ b_1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ b_{n-1} & b_{n-2} & \ldots & 0 \end{bmatrix} \begin{bmatrix} u(0) \\ u^{(1)}(0) \\ \vdots \\ u^{(n-1)}(0) \end{bmatrix}$$

Let $A(s) = s^n + a_1 s^{n-1} + \ldots + a_n$, $B(s) = b_1 s^{n-1} + b_2 s^{n-2} + \ldots + b_n$, and $V(s) = v_1 s^{n-1} + v_2 s^{n-2} + \ldots + v_n$. Equation (3) is further abbreviated as:

$$A(s)Y(s) = B(s)U(s) + V(s) \quad (4)$$

In most of the cases, the initial conditions are all zeros. If some of the initial conditions are unknown, they can be considered as the unknown parameters to be estimated. Dividing both sides of equation (3) by s^n and rearranging it, we get

$$Y(s) = -a_1 s^{-1}Y(s) - \ldots - a_n s^{-n}Y(s) + b_1 s^{-1}U(s) + \ldots + b_n s^{-n}U(s) + v_1 s^{-1} + \ldots + v_n s^{-n} \quad (5)$$

Taking the inverse Laplace transform, we get the time domain expression as

$$y(t) = -a_1\int_0^t y(t)dt - \ldots - a_n\int_0^t\int_0^t\cdots\int_0^t y(t)dt^n + b_1\int_0^t u(t)dt + \ldots + b_n\int_0^t\int_0^t\cdots\int_0^t u(t)dt^n + v_1 + \ldots + v_n\frac{t^{n-1}}{(n-1)!} \quad (6)$$

Assume that m samples are taken at different time instances t_i (i = 1, 2, ... , m). In general, if some of the v_i's are zeros and some are not, m should be greater than or equal to the total number of the parameters to be estimated in equation (6). For convenience, we denote the number of the parameters to be estimated as ρ. Integrating equation (6) n times from 0 to t_i (i=1,2,...,m) with respect to t, we get the following matrix equation

$$\mathbf{y} = \mathbf{X}\,\theta + \xi \quad (7)$$

where $\mathbf{y}=[y(t_1), y(t_2), \ldots, y(t_m)]^T$ is the column vector of the measurements at times $t_1, t_2, \ldots, t_m$, and $\theta=[-a_1, \ldots, -a_n, b_1, \ldots, b_n, v_1, \ldots, v_n]^T$ is the column vector of the parameters to be estimated with dimension of ρ, $\mathbf{X}$ is the $m \times \rho$ coefficient matrix containing integrals of input or output, or functions of t:

$$X = \begin{bmatrix} \int_0^{t_1} y(t)dt & \ldots & \int_0^{t_1}\int_0^{t_1}\ldots\int_0^{t_1} y(t)dt^n & \int_0^{t_1} u(t)dt & \ldots & \int_0^{t_1}\int_0^{t_1}\ldots\int_0^{t_1} u(t)dt^n & 1 & \ldots & \frac{t_1^{n-1}}{(n-1)!} \\ \int_0^{t_2} y(t)dt & \ldots & \int_0^{t_2}\int_0^{t_2}\ldots\int_0^{t_2} y(t)dt^n & \int_0^{t_2} u(t)dt & \ldots & \int_0^{t_2}\int_0^{t_2}\ldots\int_0^{t_2} u(t)dt^n & 1 & \ldots & \frac{t_2^{n-1}}{(n-1)!} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \int_0^{t_m} y(t)dt & \ldots & \int_0^{t_m}\int_0^{t_m}\ldots\int_0^{t_m} y(t)dt^n & \int_0^{t_m} u(t)dt & \ldots & \int_0^{t_m}\int_0^{t_m}\ldots\int_0^{t_m} u(t)dt^n & 1 & \ldots & \frac{t_m^{n-1}}{(n-1)!} \end{bmatrix}$$

$\xi = [\xi_1, \xi_2, \ldots, \xi_m]^T$ is the column vector of the equation noise terms. These terms originate from the measurement noise in y(t). If $m = \rho$ and $\mathbf{X}^{-1}$ exists, we can solve θ uniquely from equation (7) by

$$\hat{\theta} = \mathbf{X}^{-1}\mathbf{y} \quad (8)$$

where $\hat{\theta}$ denotes the estimate of θ.

If $m > \rho$, the *linear least squares (LLS)* solution for θ is given by,

$$\hat{\theta}_{LLS} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \quad (9)$$

where $\hat{\theta}_{LLS}$ represents the estimated θ in the linear least squares sense. The estimates from equation (9) are biased, even though the direct measurement noise is white or independent at different sampling times [5,9]. In other words, the equation noise ξ is correlated or coloured. This can also be shown in the frequency domain as follows.

If we rearrange equation (4) and add a white measurement noise term to the equation, we have

$$Y(s) = \frac{B(s)}{A(s)}U(s) + \frac{V(s)}{A(s)} + E(s) \quad (10)$$

where E(s) is the Laplace transform of the white noise e(t). If we convert equation (10) back to the original format of equation (4), we have

$$A(s)Y(s) = B(s)\,U(s) + V(s) + A(s)\,E(s) \quad (11)$$

From equation (11), we can see that the equation noise A(s)E(s) is coloured, even though the direct measurement noise E(s) is white. Therefore the parameters estimated from equation (9) are biased. These parameters can be refined by the generalized linear least squares method (GLLS). The main idea behind GLLS is to whiten the equation noise using the previously estimated parameters and then to re-estimate the parameters. In other words, if the

Authorized licensed use limited to: Hong Kong Polytechnic University. Downloaded on June 25,2010 at 04:06:09 UTC from IEEE Xplore. Restrictions apply.

equation noise can be whitened, an unbiased estimation can be achieved. If a rough estimation of parameters $\hat{\theta}$ is obtained from equation (9), $\hat{A}(s)$, the estimated A(s) can be determined ($\hat{A}(s) = s^n + \hat{a}_1 s^{n-1} + \cdots + \hat{a}_n$ where $\hat{a}_1,\ldots,\hat{a}_n$ are the estimates of $a_1,\ldots,a_n$). Dividing equation (11) by $\hat{A}(s)$, we obtain

$$\frac{A(s)}{\hat{A}(s)}Y(s) = \frac{B(s)}{\hat{A}(s)}U(s) + \frac{V(s)}{\hat{A}(s)} + \frac{A(s)}{\hat{A}(s)}E(s) \qquad (12)$$

If $\hat{A}(s) \rightarrow A(s)$, the equation noise is whitened and the estimates from equation (12) would be unbiased. As the noise term will not affect the derivation of $\frac{A(s)}{\hat{A}(s)}$, $\frac{B(s)}{\hat{A}(s)}$ and $\frac{V(s)}{\hat{A}(s)}$, we can first remove it from equation (12), hence

$$Y(s) - \left[\hat{a}_1\frac{s^{n-1}}{\hat{A}(s)} + \hat{a}_2\frac{s^{n-2}}{\hat{A}(s)} + \cdots + \hat{a}_n\frac{1}{\hat{A}(s)}\right]Y(s) =$$

$$- a_1\frac{s^{n-1}}{\hat{A}(s)}Y(s) - a_2\frac{s^{n-2}}{\hat{A}(s)}Y(s) - \cdots - a_n\frac{1}{\hat{A}(s)}Y(s)$$

$$+ b_1\frac{s^{n-1}}{\hat{A}(s)}U(s) + b_2\frac{s^{n-2}}{\hat{A}(s)}U(s) + \cdots + b_n\frac{1}{\hat{A}(s)}U(s)$$

$$+ v_1\frac{s^{n-1}}{\hat{A}(s)} + v_2\frac{s^{n-2}}{\hat{A}(s)} + \cdots + v_n\frac{1}{\hat{A}(s)} \qquad (13)$$

Now, consider the estimated characteristics equation

$$\hat{A}(s) = s^n + \hat{a}_1 s^{n-1} + \cdots + \hat{a}_n \qquad (14)$$

In general, if we consider that $\hat{A}(s)$ has repeated and distinct eigenvalues (these eigenvalues are assumed to be real so that a large class of practical applications are comprised), $\hat{A}(s)$ can be written as

$$\hat{A}(s) = \left[\prod_{j=1}^{q}(s-\lambda_j)^{p_j}\right]\left[\prod_{i=1}^{n'}(s-\alpha_i)\right] \qquad (15)$$

where λ_j (j = 1, 2, ... , q) are the j-th repeated eigenvalues of $\hat{A}(s)$, p_j are the repeated of multiplicity of the j-th repeated eigenvalues, and $\sum_{j=1}^{q} p_j \le n$, $n' = n - \sum_{j=1}^{q} p_j$, α_i (i = 1, 2, ..., n') are the i-th distinct eigenvalues of $\hat{A}(s)$. $\frac{s^k}{\hat{A}(s)}$ (k = 0, 1, 2, ... , n-1) in equation (13), can be written in partial fraction form as in equation (16) (shown at the bottom of the page), where $\hat{A}'(\alpha_i) = \left.\frac{d\hat{A}(s)}{ds}\right|_{s=\alpha_i}$. The inverse Laplace transform of $\frac{s^k}{\hat{A}(s)}$, $\frac{s^k}{\hat{A}(s)}Y(s)$ and $\frac{s^k}{\hat{A}(s)}U(s)$ are given by equations (17), (18), and (19) respectively (shown at the bottom of the page), where

$$C_{p_r,r',k} = \frac{1}{(p_r - r')!}\frac{d^{p_r-r'}}{ds^{p_r-r'}}\left[(s-\lambda_r)^{p_r}\frac{s^k}{\hat{A}(s)}\right]_{s=\lambda_r} \qquad (20)$$

$$y_i(t) = \frac{y(t)\otimes e^{\alpha_i t}}{\hat{A}'(\alpha_i)} = \frac{1}{\hat{A}'(\alpha_i)}\int_0^t y(\tau)e^{\alpha_i(t-\tau)}d\tau \qquad (21)$$

$$u_i(t) = \frac{u(t)\otimes e^{\alpha_i t}}{\hat{A}'(\alpha_i)} = \frac{1}{\hat{A}'(\alpha_i)}\int_0^t u(\tau)e^{\alpha_i(t-\tau)}d\tau \qquad (22)$$

$$y_{r,r'}(t) = \frac{y(t)\otimes(t^{r'-1}e^{\lambda_r t})}{(r'-1)!} = \frac{1}{(r'-1)!}\int_0^t y(\tau)(t-\tau)^{r'-1}e^{\lambda_r(t-\tau)}d\tau \qquad (23)$$

$$u_{r,r'}(t) = \frac{u(t)\otimes(t^{r'-1}e^{\lambda_r t})}{(r'-1)!} = \frac{1}{(r'-1)!}\int_0^t u(\tau)(t-\tau)^{r'-1}e^{\lambda_r(t-\tau)}d\tau \qquad (24)$$

in which $\otimes$ is convolution integration operator, r = 1, 2,..., q; r' = 1, 2,..., p_r; k = 0, 1, 2,..., n-1; i = 1, 2,..., n' and $n' = n - \sum_{j=1}^{q} p_j$.

$$\frac{s^k}{\hat{A}(s)} = \sum_{r=1}^{q}\sum_{r'=1}^{p_r}\frac{1}{(p_r-r')!}\frac{d^{p_r-r'}}{ds^{p_r-r'}}\left[(s-\lambda_r)^{p_r}\frac{s^k}{\hat{A}(s)}\right]_{s=\lambda_r}\cdot\frac{1}{(s-\lambda_r)^{r'}} + \sum_{i=1}^{n'}\frac{\alpha_i^{\,k}}{\hat{A}'(\alpha_i)}\cdot\frac{1}{s-\alpha_i} \qquad (16)$$

$$L^{-1}\left\{\frac{s^k}{\hat{A}(s)}\right\} = \sum_{r=1}^{q}\sum_{r'=1}^{p_r}C_{p_r,r',k}\cdot\frac{t^{r'-1}e^{\lambda_r t}}{(r'-1)!} + \sum_{i=1}^{n'}\frac{\alpha_i^{\,k}e^{\alpha_i t}}{\hat{A}'(\alpha_i)} \qquad (17)$$

$$L^{-1}\left\{\frac{s^k}{\hat{A}(s)}Y(s)\right\} = \sum_{r=1}^{q}\sum_{r'=1}^{p_r}C_{p_r,r',k}\,y_{r,r'}(t) + \sum_{i=1}^{n'}\alpha_i^{\,k}y_i(t) \qquad (18)$$

$$L^{-1}\left\{\frac{s^k}{\hat{A}(s)}U(s)\right\} = \sum_{r=1}^{q}\sum_{r'=1}^{p_r}C_{p_r,r',k}\,u_{r,r'}(t) + \sum_{i=1}^{n'}\alpha_i^{\,k}u_i(t) \qquad (19)$$

Authorized licensed use limited to: Hong Kong Polytechnic University. Downloaded on June 25,2010 at 04:06:09 UTC from IEEE Xplore. Restrictions apply.

When $\hat{A}(s) \rightarrow A(s)$, the equation noise ξ will be whitened, and the *generalized linear least squares (GLLS)* estimator for the non-uniformly sampled continuous systems is given by

$$\hat{\theta}_{GLLS} = (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{r} \qquad (25)$$

where

$$\mathbf{r} = \begin{bmatrix} y(t_1) - \sum_{j=1}^{n} \hat{a}_j \left(\sum_{r=1}^{q} \sum_{r'=1}^{p_r} C_{p_r, r', n-j} y_{r,r'}(t_1) + \sum_{i=1}^{n'} \alpha_i^{n-j} y_i(t_1) \right) \\ y(t_2) - \sum_{j=1}^{n} \hat{a}_j \left(\sum_{r=1}^{q} \sum_{r'=1}^{p_r} C_{p_r, r', n-j} y_{r,r'}(t_2) + \sum_{i=1}^{n'} \alpha_i^{n-j} y_i(t_2) \right) \\ \vdots \\ y(t_m) - \sum_{j=1}^{n} \hat{a}_j \left(\sum_{r=1}^{q} \sum_{r'=1}^{p_r} C_{p_r, r', n-j} y_{r,r'}(t_m) + \sum_{i=1}^{n'} \alpha_i^{n-j} y_i(t_m) \right) \end{bmatrix}$$

and $$\mathbf{Z} = [\, \mathbf{C} \mid \mathbf{D} \mid \mathbf{G} \,]$$

in which

$$\mathbf{C} = \begin{bmatrix} \sum_{r=1}^{q} \sum_{r'=1}^{p_r} C_{p_r, r', n-1} y_{r,r'}(t_1) + \sum_{i=1}^{n'} \alpha_i^{n-1} y_i(t_1) & \cdots; & \sum_{r=1}^{q} \sum_{r'=1}^{p_r} C_{p_r, r', 0} y_{r,r'}(t_1) + \sum_{i=1}^{n'} y_i(t_1) \\ \sum_{r=1}^{q} \sum_{r'=1}^{p_r} C_{p_r, r', n-1} y_{r,r'}(t_2) + \sum_{i=1}^{n'} \alpha_i^{n-1} y_i(t_2) & \cdots; & \sum_{r=1}^{q} \sum_{r'=1}^{p_r} C_{p_r, r', 0} y_{r,r'}(t_2) + \sum_{i=1}^{n'} y_i(t_2) \\ \vdots & \vdots & \vdots \\ \sum_{r=1}^{q} \sum_{r'=1}^{p_r} C_{p_r, r', n-1} y_{r,r'}(t_m) + \sum_{i=1}^{n'} \alpha_i^{n-1} y_i(t_m) & \cdots; & \sum_{r=1}^{q} \sum_{r'=1}^{p_r} C_{p_r, r', 0} y_{r,r'}(t_m) + \sum_{i=1}^{n'} y_i(t_m) \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} \sum_{r=1}^{q} \sum_{r'=1}^{p_r} C_{p_r, r', n-1} u_{r,r'}(t_1) + \sum_{i=1}^{n'} \alpha_i^{n-1} u_i(t_1) & \cdots; & \sum_{r=1}^{q} \sum_{r'=1}^{p_r} C_{p_r, r', 0} u_{r,r'}(t_1) + \sum_{i=1}^{n'} u_i(t_1) \\ \sum_{r=1}^{q} \sum_{r'=1}^{p_r} C_{p_r, r', n-1} u_{r,r'}(t_2) + \sum_{i=1}^{n'} \alpha_i^{n-1} u_i(t_2) & \cdots; & \sum_{r=1}^{q} \sum_{r'=1}^{p_r} C_{p_r, r', 0} u_{r,r'}(t_2) + \sum_{i=1}^{n'} u_i(t_2) \\ \vdots & \vdots & \vdots \\ \sum_{r=1}^{q} \sum_{r'=1}^{p_r} C_{p_r, r', n-1} u_{r,r'}(t_m) + \sum_{i=1}^{n'} \alpha_i^{n-1} u_i(t_m) & \cdots; & \sum_{r=1}^{q} \sum_{r'=1}^{p_r} C_{p_r, r', 0} u_{r,r'}(t_m) + \sum_{i=1}^{n'} u_i(t_m) \end{bmatrix}$$

$$\mathbf{G} = \begin{bmatrix} \sum_{r=1}^{q} \sum_{r'=1}^{p_r} C_{p_r, r', n-1} \dfrac{t_1^{r'-1} e^{\lambda_i t_1}}{(r'-1)!} + \sum_{i=1}^{n'} \dfrac{\alpha_i^{n-1} e^{\alpha_i t_1}}{\hat{A}'(\alpha_i)} & \cdots; & \sum_{r=1}^{q} \sum_{r'=1}^{p_r} C_{p_r, r', 0} \dfrac{t_1^{r'-1} e^{\lambda_i t_1}}{(r'-1)!} + \sum_{i=1}^{n'} \dfrac{e^{\alpha_i t_1}}{\hat{A}'(\alpha_i)} \\ \sum_{r=1}^{q} \sum_{r'=1}^{p_r} C_{p_r, r', n-1} \dfrac{t_2^{r'-1} e^{\lambda_i t_2}}{(r'-1)!} + \sum_{i=1}^{n'} \dfrac{\alpha_i^{n-1} e^{\alpha_i t_2}}{\hat{A}'(\alpha_i)} & \cdots; & \sum_{r=1}^{q} \sum_{r'=1}^{p_r} C_{p_r, r', 0} \dfrac{t_2^{r'-1} e^{\lambda_i t_2}}{(r'-1)!} + \sum_{i=1}^{n'} \dfrac{e^{\alpha_i t_2}}{\hat{A}'(\alpha_i)} \\ \vdots & \vdots & \vdots \\ \sum_{r=1}^{q} \sum_{r'=1}^{p_r} C_{p_r, r', n-1} \dfrac{t_m^{r'-1} e^{\lambda_i t_m}}{(r'-1)!} + \sum_{i=1}^{n'} \dfrac{\alpha_i^{n-1} e^{\alpha_i t_m}}{\hat{A}'(\alpha_i)} & \cdots; & \sum_{r=1}^{q} \sum_{r'=1}^{p_r} C_{p_r, r', 0} \dfrac{t_m^{r'-1} e^{\lambda_i t_m}}{(r'-1)!} + \sum_{i=1}^{n'} \dfrac{e^{\alpha_i t_m}}{\hat{A}'(\alpha_i)} \end{bmatrix}$$

and $\hat{\theta}_{GLLS}$ represents the estimated θ in the *generalized linear least squares* senses with initial $\hat{A}(s)$ provided by equation (9). Theoretically, equation (25) is used iteratively until $\hat{A}(s) \rightarrow A(s)$. However, in practical applications, one or two iterations are sufficient to obtain satisfactory parameter estimates [5]. A case study to evaluate the GLLS algorithm for the identification of a non-uniformly sampled biomedical system with a pair of repeated eigenvalues is presented in the next section.

III. CASE STUDY

The GLLS algorithm with distinct eigenvalues has been successfully applied in non-uniformly sampled biomedical system parameter estimation and in the construction of biomedical functional images [5]. In compartment analysis, when the time series data are fitted by sum of exponential functions, it is often assumed that the eigenvalues in the compartment model are real and distinct, and the number of compartments is equal to the number of exponentials. However, some systems may contain repeated eigenvalues [3]. The repeated eigenvalues, although more difficult to be detected, may be inherent in the data from the model output. In this case, the impulse response of the system is different from those models of distinct real eigenvalues.

Figure 1 shows a two-compartment system which can have either real distinct or real repeated eigenvalues. Therefore, three possible mathematical models can be used to describe the system, (i) Model 1: $A_1(e^{\lambda_1 t} - e^{\lambda_2 t})$, 2nd order model with distinct eigenvalues, (ii) Model 2: $A_1 t e^{\lambda_1 t}$, 2nd order model with a pair of repeated eigenvalues, and (iii) Model 3: $A_1 e^{\lambda_1 t} + A_2 t e^{\lambda_1 t}$, 2nd order model with a pair of repeated eigenvalues (general expression).

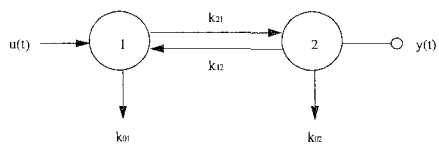

Figure 1 The two compartment model with input in compartment 1 and output in compartment 2 respectively.

Suppose $k_{21} = k_{02} = 1$, $k_{12} = k_{01} = 0$ and $u(t) = 10\delta(t)$, an impulse of magnitude 10, the impulse response of the system is $y(t) = 10te^{-t}$, which has a pair of repeated eigenvalues $\lambda_1=\lambda_2=-1$. In other words, Model 2 is the right model for the system. For the noise free data generated by $y(t) = 10te^{-t}$, we can easily use the GLLS algorithm to validate that Model 2 is the right model. To generate the noisy data, we added zero-mean white Gaussian noise with 5% coefficients of variation (CV) to y(t), the model output. Model 1, Model 2 and Model 3 were then used to fit the noisy data using the GLLS algorithm respectively. The simulation study was performed using the PV-WAVE package.

Numerous statistical tests can be used to compare the results of parameters estimation for determining the best fit model. In this paper, we use the analysis of weighted residuals sum of squares (WRSS), parameters estimate asymptotic coefficients of variation (CVs), the parameter correlation matrix (R), Akaike Information Criterion (AIC) [1] and Schwarz Criterion (SC) [8] to determine which model is the best.

The results of the parameter estimation are summarized in Table 1. As seen in Table 1, the

Authorized licensed use limited to: Hong Kong Polytechnic University. Downloaded on June 25,2010 at 04:06:09 UTC from IEEE Xplore. Restrictions apply.

weighted residual sum of squares (WRSS) of the three models are very close, other criteria should be used to determine the best fit model. Compare Model 1 with Model 2, although the WRSS of Model 1 is smaller than that of Model 2, the parameter estimates in Model 1 have the largest asymptotic CVs whereas the asymptotic CVs of the parameters in Model 2 are much smaller. Moreover, the AICs and SCs of Model 1 are larger than those of Model 2, which are also in favour to Model 2. Furthermore, the parameter estimates in Model 1 are strongly correlated as shown in the correlation matrix $\mathbf{R}_{\text{Model 1}}$, whereas the correlation matrix of Model 2, $\mathbf{R}_{\text{Model 2}}$, shows that the parameter estimates in Model 2 are less correlated.

$$\mathbf{R}_{\text{Model 1}} = \begin{array}{c} \\ A_1 \\ \lambda_1 \\ \lambda_2 \end{array} \begin{array}{c} \begin{array}{ccc} A_1 & \lambda_1 & \lambda_2 \end{array} \\ \begin{bmatrix} 1.0000 & -0.9869 & 0.9897 \\ -0.9869 & 1.0000 & -0.9615 \\ 0.9897 & -0.9615 & 1.0000 \end{bmatrix} \end{array}$$

$$\mathbf{R}_{\text{Model 2}} = \begin{array}{c} \\ A_1 \\ \lambda_1 \end{array} \begin{array}{c} \begin{array}{cc} A_1 & \lambda_1 \end{array} \\ \begin{bmatrix} 1.0000 & -0.6398 \\ -0.6398 & 1.0000 \end{bmatrix} \end{array}$$

Based on the above arguments, it can be seen that Model 2 is better than Model 1 and we can reject Model 1 from our analysis. Now, we compare Model 2 with Model 3. Both models contain repeated eigenvalues and Model 3 being the general expression. The correlation matrix of Model 3 is given by

$$\mathbf{R}_{\text{Model 3}} = \begin{array}{c} \\ A_1 \\ \lambda_1 \\ A_2 \end{array} \begin{array}{c} \begin{array}{ccc} A_1 & \lambda_1 & A_2 \end{array} \\ \begin{bmatrix} 1.0000 & 0.1183 & -0.1830 \\ 0.1183 & 1.0000 & -0.6563 \\ -0.1830 & -0.6563 & 1.0000 \end{bmatrix} \end{array}$$

From Table 1, it can be seen that the WRSS of Model 2 is less than that of Model 3. Moreover, comparison of the sub-matrix of $\mathbf{R}_{\text{Model 3}}$ (the resultant matrix of $\mathbf{R}_{\text{Model 3}}$ after removing the first row and the first column of $\mathbf{R}_{\text{Model 3}}$) and the correlation matrix of Model 2 ($\mathbf{R}_{\text{Model 2}}$) also shows that Model 2 is better than Model 3 since the parameters in Model 2 are less correlated than those of Model 3. In addition, one of the parameters (A_1) in Model 3 has a very large asymptotic CV (11.10%) and the asymptotic CVs of the other two parameters (A_2 and λ_1) are slightly larger than those corresponding to the parameters A_1 and λ_1 in Model 2. The above comparisons suggest that Model 3 is not as suitable as Model 2 to describe the system. To further support of our arguments, we can consider the AICs and SCs of both models. As shown in Table 1, the AICs and SCs of Model 2 are smaller than those of Model 3, which in turn implied that Model 3 is over-parameterized to describe the system. In other words, Model 2 is better than Model 3.

Comparisons of the two rejected models (Model 1 and Model 3) also suggest Model 3 is still better than Model 1 since the correlation coefficients and the asymptotic CVs of the parameters in Model 3 are much smaller than those of Model 1.

Based on the above comparative arguments, we conclude that the best model to describe the system is Model 2 which is the exact form of the impulse response of the system that generated the data.

IV. CONCLUSION

In this paper, a fast algorithm is presented, which can estimate continuous-time model parameters directly without the need of providing initial parameter values. Moreover, the algorithm can produce unbiased parameter estimates and it requires very little computing time. The case study presented demonstrates the reliability of the algorithm. With this algorithm, we can provide more choices for system identification and can find the best suitable model for the system from the non-uniformly sampled noisy data.

Table 1 Parameter estimates and statistical results of the three possible models

	Model 1			Model 2			Model 3		
Parameter	Estimator	SD	CV(%)	Estimator	SD	CV(%)	Estimator	SD	CV(%)
A_1	28.09750	±4.02653	±15.18	10.16950	±0.12733	±1.25	0.00859	±0.00095	±11.10
λ_1	-0.85642	±0.01612	±1.89	-1.03248	±0.00325	±0.31	-1.03814	±0.00331	±0.32
A_2	-	-	-	-	-	-	10.21165	±0.13371	±1.31
λ_2	-1.21986	±0.04274	±3.50	-	-	-	-	-	-
Summary Statistics									
WRSS	57.53785			57.74197			59.99215		
df[a]	24			25			24		
VR[b]	2.39741			2.30969			2.49967		
Case (i): Weights chosen exactly equal to 1/(error variance)									
AIC	63.53785			61.74197			65.99215		
SC	67.42536			64.33364			69.87966		
Case (ii): Error variance only known to within a proportional constant K									
AIC	115.41596			113.51157			116.54377		
SC	119.30347			116.10325			120.43128		

[a]Degree of Freedom = number of data point(N) - number of parameters(P) + number of constraints
[b]Variance Ratio = WRSS/df

Authorized licensed use limited to: Hong Kong Polytechnic University. Downloaded on June 25,2010 at 04:06:09 UTC from IEEE Xplore. Restrictions apply.

ACKNOWLEDGEMENT

We would like to acknowledge the support from the Australian Research Council (ARC) and the Research Committee of the Hong Kong Polytechnic University.

REFERENCES

[1] Akaike H., "A Bayesian analysis of the minimum AIC procedure", *Ann. Inst. Stat. Math.*, vol. 30, pp. 9-14, 1978.

[2] Åström K.J., "Maximum-likelihood and prediction error methods", *Automatica*, vol.16, pp.551-557, 1980.

[3] Fagarasan J.T. and DiStefano J.J. III, "Hidden pools, hidden modes and visible repeated eigenvalues in compartmental models", *Math. Biosci.*, vol. 82, pp.87-113, 1986.

[4] Feng D., Ho D., Chen K., Wu L., Wang J., Liu R., and Yeh S., "An Evaluation of the Algorithms for Determining Local Cerebral Metabolic Rates of Glucose Using Positron Emission Tomography Dynamic Data", *IEEE Trans. Med. Imag.*, vol. 14, No. 4, pp. 697-710, 1995.

[5] Feng D., Huang S.C., Wang Z., and Ho D., "An Unbiased Parametric Imaging Algorithm for Non-Uniformly Sampled Biomedical System Parameter Estimation", *IEEE Trans. Med. Imag.*, accepted, 2/1996.

[6] Hsia T.C., "System Identification: Least Squares", *Lexington Books, D.C. Heath and Company*, 1977.

[7] Mori F. and DiStefano J.J., III, "Optimal Non-uniform Sampling Interval and Test Input Design for Identification of Physiological Systems from Very Limited Data", *IEEE Trans. Automatic Control*, vol. 24, pp. 893-899, 1979.

[8] Schwarz G., "Estimation of the dimension of a model", *Ann. Stat.*, vol. 6, pp.461-464, 1978.

[9] Wong K.P., Feng D., and Siu W.C., "Generalized Linear Least Squares Algorithm for Non-uniformly Sampled Biomedical System Identification With Possible Repeated Eigenvalues", *Submitted to International Journal of Systems Science*, 5/1996.

[10] Yamamoto Y.L., Thompson C.J., Meyer E.E., Robertson J.S., and Feindel W., "Dynamic Positron Emission Tomography for study of Cerebral Hemodynamics in a Cross Section of Head Using Positron Emitting ^{68}Ga-EDTA and ^{77}Kr", *J. Comput. Assist. Tomogr.*, vol.1, pp.43-56, 1977.

Authorized licensed use limited to: Hong Kong Polytechnic University. Downloaded on June 25,2010 at 04:06:09 UTC from IEEE Xplore. Restrictions apply.