Serial Number 09/536,488

Filing Date 27 March 2000

Inventor Francis J. O'Brien, Jr.
Chung T. Nguyen
Bruce J. Bates

NOTICE

The above identified patent application is available for licensing. Requests for information should be addressed to:

OFFICE OF NAVAL RESEARCH
DEPARTMENT OF THE NAVY
CODE 00CC
ARLINGTON VA 22217-5660

Attorney Docket No. 79236

ENHANCED MODEL IDENTIFICATION IN SIGNAL PROCESSING USING ARBITRARY EXPONENTIAL FUNCTIONS

STATEMENT OF GOVERNMENT INTEREST

The invention described herein may be manufactured and used by or for the Government of the United States of America for governmental purposes without the payment of any royalties thereon or therefore.

BACKGROUND OF THE INVENTION

(1) Field of the Invention

This invention relates to a computer-aided method for signal or data processing and more particularly to a method for finding, by means of nonlinear regression analyses, a Probability Density Function (PDF) for arbitrary exponential functions falling into one of four classes, the underlying probability density function and data structures conforming to the exponential model.

(2) Description of the Prior Art

Of the many continuous probability distributions encountered in signal processing, a good number are distinguished by the fact that they are derived from exponential functions on the time interval of 0 to ∞, e.g., failure rate distributions, Poisson processes, Chi-Square, gamma, Rayleigh, Weibull, Maxwell and others involving exponential functions. Such an exponential function is also used in O'Brien et al. (U.S. Patent No. 5,537,368) to generate a corrected data stream from the raw data stream of a sensor.

Occasionally modeling involves functions for which the probability density function (PDF) and its moments need to be derived de novo. Often times, research scientists and engineers are confronted with modeling a random variable x when the probability density function (PDF) is unknown. It may be known that the variable can be reasonably well approximated by a gamma density. Then solving a problem under the assumption that x has a gamma density will provide some insight into the true situation. This approach is all the more reasonable since many probability distributions are related to the gamma function. However, deriving the PDF and its statistical moments using the standard approach involving moment generating functions (MGF) and complex-variable characteristic functions is difficult and somewhat impractical to implement in applied research settings. The complexity of current methods for constructing the PDF and MGF limits the class of models used for analyzing correlated data structures. O'Brien et al. (U.S. Patent No. 5,784,297) provided a method for finding a probability density function (PDF) and its statistical moments for an arbitrary exponential function of the form $g(x) = \alpha x^m e^{-\beta x^n}$, $0 < x < \infty$, where α, β, $n > 0$, $m > -1$ are real constants in one-dimensional distributions and $g(x_1, x_2, \ldots, x_l)$ in the hyperplane. However, the method in the '297 patent is based on a single probability model within the domain $0 \rightarrow \infty$, thus limiting its application.

SUMMARY OF THE INVENTION

Accordingly, it is a general purpose and object of the present invention to provide a computer-aided method for determining density and moment functions for a useful class of exponential functions in signal processing.

Another object of the present invention is to provide a method for constructing the PDF and MGF which offers the possibility of constructing the PDF and MGF for a larger class of such functions.

A still further object is to enhance standard assumptions about the structure of error or disturbance terms by including a larger class of models to choose from.

These objects are provided with the present invention by a simple substitution method for finding a probability density function (PDF) and its statistical moments for a chosen one of four newly derived probability models for an arbitrary exponential function of the forms $g(x) = \alpha x^m e^{-\beta x^n}$, $-\infty < x < \infty$; $g(x) = \alpha x^m e^{-\beta x^n}$, $0 \le x < \infty$; $g(x) = \alpha\left(\frac{x-a}{b}\right)^m e^{-\beta\left(\frac{x-a}{b}\right)^n}$, $-\infty < x < \infty$; and $g(x) = \alpha\left(\frac{x-a}{b}\right)^m e^{-\beta\left(\frac{x-a}{b}\right)^n}$, $0 \le x < \infty$. The model chosen will depend on the domain of the data and whether information on the parameters *a* and *b* exists. These parameters may typically be the mean or average of the data and the standard deviation, respectively. For example, it may be known that the signal of interest within the data being processed has a domain from $-\infty \rightarrow \infty$ and a typical mean and standard deviation. Thus a model of the third form would be used.

Once the model is chosen, computer implemented non-linear regression analyses are performed on the data distribution to determine the solution set $S_n(\alpha_n, m_n, \beta_n, n)$ beginning with $n = 1$. A root-mean-square (RMS) is calculated and recorded for each order of n until the regression analyses produce associated RMS values that are not changing in value appreciably. The basis function is reconstructed from the estimates in the final solution set and a PDF for the basis function is obtained utilizing methods well known in the art. The MGF, which characterizes any statistical moment of the distribution, is obtained using a novel function derived by the inventors and the mean and variance are obtained in standard fashion. Once the parameters α, β, m and n have been determined for a set of data measurements through the system identification modeling, the PDF-based mean and variance are determinable, and simple binary hypotheses may be tested.

By the inclusion of four newly derived models, the method of the present invention provides a choice of models from a larger and more useful class of exponential functions covering the full domain ($-\infty \rightarrow \infty$). The method of the present invention further provides enhanced standard assumptions about the structure of error or disturbance terms by the use of additional variables such as mean and standard deviation parameters.

BRIEF DESCRIPTION OF THE DRAWINGS

A more complete understanding of the invention and many of the attendant advantages thereto will be readily appreciated as the same becomes better understood by reference to the following detailed description when considered in conjunction with the accompanying drawings wherein corresponding reference characters indicate corresponding parts throughout the several views of the drawings and wherein:

FIG. 1 shows a set of data points appearing to conform to a negative exponential (or decay) function; and

FIG. 2 is a flow chart of the steps used to identify and characterize the function of FIG. 1.

DESCRIPTION OF THE PREFERRED EMBODIMENT

Referring to FIG. 1, there is shown a plot of a measured time series. Each point "p" corresponds to a measurement taken by a sensor at a certain time. Such a plot may correspond to the output signals from a sonar sensor. In order to analyze the signals, it is necessary to identify a frequency function, $g(x)$, which models or characterizes the data structure. Referring now also to FIG. 2, there is shown a flow chart of the steps used to obtain $g(x)$. The data are first obtained in step 10 and then the measured time series is plotted at step 12, as has been done in FIG. 1 to determine, at step 14, if an exponential decay function is reasonable. It can be seen from FIG. 1 that as time increases, the measured value decreases. Such a decreasing function over time may correspond to a negative exponential (or decay) function. Since the data points of FIG. 1 appear to correspond to an exponential decay function, step 14 transfers control to step 16. If an exponential decay function was not reasonable, step 14 would exit from the method and a different model approach would be required to analyze the data.

Step 16 chooses a model to characterize the data, depending on the domain of the data and whether certain parameter information, such as mean and standard deviation, is known about the data. As can be seen from the grid in step 16, Model I and Model III are applicable where the domain is $-\infty \rightarrow \infty$; Model II and Model IV are applicable for domains of $0 \rightarrow \infty$; Model I and

Model II are applicable when no parameter information is available; and Models III and Model IV are applicable when parameter information is available.

For FIG. 1, $g(x)$ is shown as a sloping-down arc and is assumed to be an optimum least-squares solution derived for the discrete time series data points "p". It is noted that the domain for the data is $0 \rightarrow \infty$. In addition, for the example of FIG. 1, it will be assumed that no parameter information is available, thus Model II is appropriate. Let the function $g(x)$ for Model II be denoted by

$$g(x) = \alpha x^{m} e^{-\beta x^{n}},\ 0 < x < \infty \qquad (1)$$

where α, β, $n > 0$, $m > -1$ are real constants. The function $g(x)$ is obtained in the standard manner for exponential functions. First take the natural logarithm of the modeling basis, or exponential function of Equation (1):

$$\log[g(x)] = \log\alpha + m\log x - \beta x^{n}. \qquad (2)$$

Because the term βx^{n} in nonlinearizable, a nonlinear approach must be taken. This approach consists of performing regression analysis on $g(x)$, with the nonlinear parameter n set to a specific integer value n = 1, 2, 3, etc. Each $[x,y]$ observation in the sample is indexed with the subscript i, where i runs from 1 to p, p being the total number of data points. Nonlinear regression analyses are performed on the data distribution based on the least squares minimization criterion, stated as follows:

$$S_n(\alpha,m,\beta,n)=\sum_{i=1}^{p}\left[\log y_i-\log\alpha-m\log x_i+\beta x_i^n\right]^2\rightarrow \min \quad (3)$$

where α, β, *m* and *n* are real-valued constants that we seek to identify through classical least squares regression analyses. Step 18 begins the regression analyses by first setting *n* = 1.

In step 20 the regression is performed using the ordinary least squares (OLS) algorithm, well known to those skilled in the art, beginning with the parameter *n* = 1 from step 18. The result is the first solution set, $S_1(\alpha_1,m_1,\beta_1,1)$, with parameters α, β and *m* estimated. The solution set is recorded and stored. In step 22 a measure of the adequacy of solution set $S_1(\alpha_1,m_1,\beta_1,1)$ is obtained and, in standard engineering fashion, the root-mean-square (RMS) statistic is calculated and recorded. Since *n* = 1, step 24 passes control to step 26 which increments parameter *n* and returns to steps 20 and 22. With *n* now having a value of 2, step 20 calculates a new solution set $S_2(\alpha_2,m_2,\beta_2,2)$, and step 22 calculates a new RMS statistic. Since *n* is now greater than 1, step 24 passes control to step 28 which tests for convergence of the associated RMS statistics. If the associated RMS values are changing in value more than a chosen convergence threshold, step 26 is repeated so as to increment *n* and steps 20 through 28 are repeated until at some value *n* = *k*, the associated RMS values are not changing in value by more than the chosen convergence threshold. The solution set $S_k(\alpha_k,m_k,\beta_k,k)$ is then called the "optimum" solution. Step 30 obtains the basis function which is the exponential function reconstructed from the estimates in the final solution set $S_k(\alpha_k, m_k, \beta_k, k)$. The basis function will be identified as follows:

$$g(x) = \alpha x^m e^{-\beta x^k}, \tag{4}$$

obtained from the parameter estimate set $S_k(\alpha_k, m_k, \beta_k, k)$, the empirical least squares solution to the data generated in the time series. In step 32, the probability density function (PDF), a concept well known to those in the art, is obtained from the basis function. The moment generating function (MGF) is then obtained in step 34 and the mean and variance are obtained from the MGF in step 36.

The mathematics involved in obtaining the PDF and MGF are quite complex. The method of the present invention utilizes two, independently derived improper definite integrals based on the general exponential integral formula in F.J. O'Brien, S.E. Hammel and C.T. Nguyen, "The Moi Formula for Improper Exponential Definite Integrals," *Perceptual and Motor Skills*, 79, 1994, pp. 1123-1127, and presented in F.J. O'Brien, S.E. Hammel and C.T. Nguyen, "The Moi Formula," accepted in I.S. Gradshteyn and I.M. Ryzhik, *Table of Integrals, Series and Products*, Academic Press (New York 1994). These two new integrals take the forms:

$$\int_{-\infty}^{\infty} \alpha x^m e^{-\beta x^n} dx = \begin{cases} 2\dfrac{\alpha\Gamma(\gamma)}{n\beta^\gamma}, \text{for even function on } -\infty < x < \infty \\ \quad \text{where } \gamma = \dfrac{m+1}{n} > 0, m \geq 0, \alpha, \beta, n > 0 \\ 0, \text{ odd function} \end{cases} \quad ; \text{ and} \tag{5}$$

$$\int_0^\infty \alpha x^m e^{-\beta x^n} dx = \frac{\alpha\Gamma(\gamma)}{n\beta^\gamma} \text{ for } 0 \le x < \infty \text{ where } \gamma = \frac{m+1}{n} > 0, m > -1, \alpha, \beta, n > 0, \quad (6)$$

where $\Gamma(\gamma)$ represents the standard gamma function in Equations (5) and (6). It will be appreciated that many real-valued one-dimensional and, by extension, multidimensional exponential functions conform to those two integrals above including functions (comprising the integrands) which must first be manipulated algebraically and/or analytically by means of change of variable, substitution, binomial expansion, completing the square or first-order differential equation analysis manipulation, inter alia. The first model, Model I, involves the integral of (5) and takes the form:

$$\int_{-\infty}^{\infty} \alpha x^m e^{-\beta x^n} dx, -\infty < x < \infty. \quad (7)$$

The second model, Model II, is based on Equation (6):

$$\int_0^\infty \alpha x^m e^{-\beta x^n} dx \quad 0 \le x < \infty. \quad (8)$$

The general integral of Model III is based on Equation (5):

$$\int_{-\infty}^{\infty} \alpha \left(\frac{x-a}{b}\right)^m e^{-\beta\left(\frac{x-a}{b}\right)^n} dx, \quad -\infty < x < \infty. \quad (9)$$

The last model, Model IV, arises less often than others, but it is a valid expression of probability models to be considered. The integral of interest is, based on (2):

$$\int_0^{\infty} \alpha\left(\frac{x-a}{b}\right)^m e^{-\beta\left(\frac{x-a}{b}\right)^n} dx, \quad 0 \le x < \infty . \qquad (10)$$

In both Equations (9) and (10), *a* and *b* are some constants selected by the practitioner and represent parameters of the data for which the user of the method has some knowledge. As previously noted, these parameters may typically be the mean and standard deviation of the data. Solving Equations (9) and (10) involves a change of variable such that $s = \frac{x-a}{b}$ and $ds/dx = b^{-1}$. Then it can readily be seen that Equation (9) falls under Equation (5) and Equation (10) falls under Equation (6).

Use of the Equations (7) through (10) will simplify the mathematics involved in deriving PDF's and moments for a useful class of continuous functions. Table 1 lists twelve frequently encountered continuous probability density functions (PDF's) taken from standard sources such as P.J. Hoel, et al., *Introduction to Probability Theory,* Houghton-Mifflin (Boston, 1971) and Abramowitz, M., Stegun, I.A., chapter 26, *Handbook of Mathematical Functions,* Washington, D.C. (Government Printing Office 1964). Each of those densities can be classified into one of the four probability models above and expressed in terms of the seven parameters, *a*, *b*, α, *m*, *n*, β and γ of the above four classes of exponential functions. Each of the densities in Table 1 is distinguished by the fact that when integrated over its appropriate interval, 0 (or -∞) to ∞, each is equal to 1, the definition of a PDF.

TABLE 1

Selected Univariate Densities Based on Exponential Functions

Density Name	Domain	Probability density function, $f(x)$	Parameter Restriction
Exponential	$0 \le x < \infty$	$\lambda e^{-\lambda x}$	$0 < \lambda < \infty$
Gamma	$0 \le x < \infty$	$\frac{\lambda^p}{\Gamma(p)} x^{p-1} e^{-\lambda x}$	$0 < p < \infty$ $0 < \lambda < \infty$
Chi-Square	$0 \le x < \infty$	$\frac{x^{\frac{\nu}{2}-1}}{\Gamma\left(\frac{\nu}{2}\right) 2^{\frac{\nu}{2}}} e^{-\frac{x}{2}}$	$0 < n < \infty$
Rayleigh	$0 \le x < \infty$	$2axe^{-ax^2}$	$0 < a < \infty$
Gamma-Poisson	$0 \le x < \infty$	$\frac{d(c\lambda)^m}{\Gamma(m)} x^{md-1} e^{-c\lambda x^d}$	$0 < m < \infty$ $0 < \lambda < \infty$ $0 < c < \infty$ $2 \le d < \infty$
Weibull	$0 \le x < \infty$	$abx^{b-1} e^{-ax^b}$	$0 < a < \infty$ $0 < b < \infty$
Maxwell	$0 \le x < \infty$	$\sqrt{2/\pi}\, x^2 e^{-\frac{1}{2}x^2}$	none
Error Function	$-\infty < x < \infty$	$\frac{h}{\sqrt{\pi}} e^{-h^2 x^2}$	$0 < h < \infty$

TABLE 1 (cont.)

Density Name	Domain	Probability density function, $f(x)$	Parameter Restriction
Normal (Standardized)	$-\infty < x < \infty$	$\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$	none
Normal (Non Standardized)	$-\infty < x < \infty$	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$	$-\infty < \mu < \infty$ $0 < \sigma < \infty$
Laplace	$-\infty < x < \infty$	$\frac{1}{2b} e^{-\left\lvert\frac{x-a}{b}\right\rvert}$	$-\infty < a < \infty$ $0 < b < \infty$
Pearson (Type III)	$a \leq x < \infty$	$\frac{1}{b\Gamma(p)}\left(\frac{x-a}{b}\right)^{p-1} e^{-\left(\frac{x-a}{b}\right)}$	$-\infty < a < \infty$ $0 < b < \infty$ $0 < p < \infty$

Equations (5) and (6) are used in step 32 to find a one-dimensional probability density function (PDF) for one of the four classes (Models I through IV) of exponential functions chosen at step 16. Abramowitz provides a summary of the mathematical properties comprising a PDF. For the purposes of this method, a PDF *f(x)* is assumed to be a real-valued non negative function. A PDF for Model I and Model III will also be an even function. In Equation (5), *m* and *n* are even numbers or fractions with even numerators. Continuing with the example of Model II used previously, then for any function corresponding to Equation (1), the PDF f(x) is given by standard integral calculus techniques applied to bounded, improper exponential definite integrals:

$$f(x) = \frac{n\beta^{\gamma}}{\Gamma(\gamma)} x^m e^{-\beta x^n} \tag{11}$$
$$f(x) \geq 0$$
$$\int_0^{\infty} f(x)dx = 1\,,$$

where f(x) will denote the PDF of an arbitrary distribution. Using similar reasoning, it can be demonstrated that for the other three probability models, the corresponding PDF's are as given in Table 2.

TABLE 2.

Density Functions

Model	Domain	Integral Equation	Normalizing Constant *c*	PDF *f(x)*
I	$-\infty<x<\infty$	$\int_{-\infty}^{\infty} \alpha x^m e^{-\beta x^n} dx$	$\frac{n\beta^{\gamma}}{2\alpha\Gamma(\gamma)}$	$\frac{n\beta^{\gamma}}{2\Gamma(\gamma)} x^m e^{-\beta x^n}$
II	$0\leq x<\infty$	$\int_{0}^{\infty} \alpha x^m e^{-\beta x^n} dx$	$\frac{n\beta^{\gamma}}{\alpha\Gamma(\gamma)}$	$\frac{n\beta^{\gamma}}{\Gamma(\gamma)} x^m e^{-\beta x^n}$
III	$-\infty<x<\infty$	$\int_{-\infty}^{\infty} \alpha \left(\frac{x-a}{b}\right)^m e^{-\beta\left(\frac{x-a}{b}\right)^n} dx$	$\frac{n\beta^{\gamma}}{2\alpha b\Gamma(\gamma)}$	$\frac{n\beta^{\gamma}}{2b\Gamma(\gamma)} \left(\frac{x-a}{b}\right)^m e^{-\beta\left(\frac{x-a}{b}\right)^n}$
IV	$0\leq x<\infty$	$\int_{0}^{\infty} \alpha \left(\frac{x-a}{b}\right)^m e^{-\beta\left(\frac{x-a}{b}\right)^n} dx$	$\frac{n\beta^{\gamma}}{\alpha b\Gamma(\gamma)}$	$\frac{n\beta^{\gamma}}{b\Gamma(\gamma)} \left(\frac{x-a}{b}\right)^m e^{-\beta\left(\frac{x-a}{b}\right)^n}$

The following parameter restrictions apply to the models:

Model I (even function PDF)
$$\gamma = \frac{m+1}{n} > 0$$
$$m \geq 0,\ n, \alpha, \beta > 0$$

Model II
$$\gamma = \frac{m+1}{n} > 0$$
$$m > -1$$
$$\alpha, \beta, n > 0$$

Model III (even function PDF)
$$\gamma = \frac{m+1}{n} > 0$$
$$m \geq 0,\ n, \alpha, \beta, b > 0$$
$$-\infty < a < \infty$$

Model IV
$$\gamma = \frac{m+1}{n} > 0$$
$$m > -1$$
$$\alpha, \beta, n, b > 0$$
$$-\infty < a < \infty$$

Once the PDF is obtained, the moment generating function (MGF), which characterizes any statistical moment of the distribution, is obtained in step 34. The moments of a probability density function are important for several reasons. The first moment corresponds to the mean of the distribution, and the second moment allows a calculation of the dispersion or variance of the distribution as indicated in step 36. The mean and variance may then be used in the central limit theorem or normal approximation formula for purposes of hypothesis testing. Additional moment-based relations such as skewness and kurtosis coefficients can also be calculated.

The moment generating functions (MGF) for the four models are shown in Tables 3a (moments about the origin) and 3b (moments about the mean:

TABLE 3a.

Moments About Origin Function

Model	Moments Function, EX^j	Notes
I	$\dfrac{\beta^{-j/n}\Gamma(\gamma + j/n)}{\Gamma(\gamma)}$	$EX^j = 0$, j odd ($j \geq 1$)
II.	$\dfrac{\beta^{-j/n}\Gamma(\gamma + j/n)}{\Gamma(\gamma)}$	
III	$\dfrac{j!}{\Gamma(\gamma)}\sum_{k=0}^{2k=j}\dfrac{a^{j-2k}b^{2k}\beta^{-2k/n}\Gamma\left(\gamma + \frac{2k}{n}\right)}{(2k)!(j-2k)!}$	$EX^j = 0$, j odd ($j \geq 3$)
IV	$\dfrac{j!}{\Gamma(\gamma)}\sum_{k=0}^{j}\dfrac{a^{j-k}b^{k}\beta^{-k/n}\Gamma\left(\gamma + \frac{k}{n}\right)}{k!(j-k)!}$	

TABLE 3b.

Moments About Mean Functions

Model	Moments Function, $E(X-\mu)^j$	Notes
I	$\dfrac{\beta^{-j/n}\Gamma(\gamma+j/n)}{\Gamma(\gamma)}$	$E(X-\mu)^j = 0$, j odd
II	$\dfrac{j!\beta^{-j/n}}{\Gamma(\gamma)}\sum_{k=0}^{j}(-1)^{j-k}\dfrac{\left[\dfrac{\Gamma\left(\gamma+\frac{1}{n}\right)}{\Gamma(\gamma)}\right]^{j-k}\Gamma\left(\gamma+\frac{k}{n}\right)}{k!(j-k)!}$	
III	$\dfrac{b^j\beta^{-j/n}\Gamma(\gamma+j/n)}{\Gamma(\gamma)}$	$E(X-\mu)^j = 0$, j odd ($j \geq 3$)
IV	$\dfrac{j!b^j\beta^{-j/n}}{\Gamma(\gamma)}\sum_{k=0}^{j}(-1)^{j-k}\dfrac{\left[\dfrac{\Gamma\left(\gamma+\frac{1}{n}\right)}{\Gamma(\gamma)}\right]^{j-k}\Gamma\left(\gamma+\frac{k}{n}\right)}{k!(j-k)!}$	

The mean and the variance are obtained in step 36 in standard fashion well known in the art. The mean is defined as $\mu = E(x)$, and the variance is defined as $\sigma^2 = E(x^2) - \mu^2$. The mean and variance for the four models are shown in Table 4:

TABLE 4.

Means and Variances

Model	Mean, μ	Variance, σ^2
I	0	$\beta^{-2/n}\dfrac{\Gamma\left(\gamma+\frac{2}{n}\right)}{\Gamma(\gamma)}$
II	$\dfrac{\beta^{-1/n}\Gamma(\gamma+1/n)}{\Gamma(\gamma)}$	$\beta^{-2/n}\left\{\dfrac{\Gamma\left(\gamma+\frac{2}{n}\right)}{\Gamma(\gamma)}-\left[\dfrac{\Gamma\left(\gamma+\frac{1}{n}\right)}{\Gamma(\gamma)}\right]^2\right\}$
III	a	$b^2\beta^{-2/n}\dfrac{\Gamma\left(\gamma+\frac{2}{n}\right)}{\Gamma(\gamma)}$
IV	$a+b\beta^{-1/n}\dfrac{\Gamma\left(\gamma+1/n\right)}{\Gamma(\gamma)}$	$b^2\beta^{-2/n}\left\{\dfrac{\Gamma\left(\gamma+\frac{2}{n}\right)}{\Gamma(\gamma)}-\left[\dfrac{\Gamma\left(\gamma+\frac{1}{n}\right)}{\Gamma(\gamma)}\right]^2\right\}$

The mean and variance may then be used in the central limit theorem or normal approximation formula for purposes of hypothesis testing in step 38, the primary use of the method being to test statistically hypotheses about the behavior of such functional forms once the empirical least squares methods have identified an applicable model derived from actual measurements. The central limit theorem or normal approximation formulas are typically of interest to those skilled in the art for evaluation of simple hypotheses. *Chebychev's Theorem,* which gives the probability of deviation from a mean regardless of the distribution, may also be of interest. The results of the hypotheses testing of step 38 can then be used to direct additional data gathering at step 40.

Confirmatory calculations, based on the entire method for the known probability distributions of Table 1, substantiate the correctness of the model calculations in that they agree with well known published results from statistical literature. In addition, the densities of Table 1 can be classified into the four models as follows:

Model I - Error Function and Normal (Standardized);

Model II - Exponential, Gamma, Chi-Square, Rayleigh, Gamma-Poisson, Weibull and Maxwell;

Model III - Normal (non-Standardized) and Laplace; and

Model IV - Pearson (Type III).

It can be seen from the iterative nature of the OLS algorithm being used that the method is suitable for implementation on computer 50, shown encompassing steps 18 through 40. Depending on the nature of the sensors being used to generate the data and on the nature of the data itself, the functions of computer 50 may include obtaining the data at step 10 through choosing a model at step 16, as shown by enlarged computer portion 50a. Alternately, steps 10 through 16 may be performed by a user and the results input into computer 50 to perform the regression analysis.

It is to be noted that the form of Equation (3) and the basis function, Equation (4) vary corresponding to the model chosen at step 16. Equation (3) is used for Model I and II, the limits being $-\infty < (x_i, y_i) < \infty$ for Model I and $0 \leq (x_i, y_i) < \infty$ for Model II. For Models III and IV, a substitution is made into Equation (3) yield:

$$S_n(\alpha, m, \beta, n) = \sum_{i=1}^{p}\left[\log y_i - \log\alpha - m\log\left(\frac{x_i - a}{b}\right) + \beta\left(\frac{x_i - a}{b}\right)^n\right]^2 \to \min, \qquad (12)$$

Models III and IV having limits corresponding to Models I and II, respectively. The basis functions for Model I are the same as Equation (4) with limits as shown above. As with Equation (3), the basis function for Models III and IV is also obtained by substitution and having the same limits as above:

$$g(x) = \alpha\left(\frac{x-a}{b}\right)^m e^{-\beta\left(\frac{x-a}{b}\right)^n}. \qquad (4)$$

What has thus been described is a method which offers a general solution for determining density and moment functions for a useful class of exponential functions in signal processing. The present method offers the possibility of constructing the PDF and MGF for a much larger class of such functions than the standard distributions, such as those listed in Table 1. Moreover, standard assumptions about the structure of error or disturbance terms can be enhanced by including a larger class of models to choose from. Many alternative or additional approaches can be introduced into the method disclosed. For example, the regression analyses can be performed with the key parameter *n* set to noninteger values. Also, measures other than the standard RMS statistic, such as the normalized "squared statistical correlation coefficient", can be used to judge the degree of fit to the distribution. Further, many real-valued, one-dimensional and, by extension, multidimensional exponential functions conform to the two integrals Equations (5) and (6) but which must first be manipulated algebraically and/or analytically by means of change of variable, substitution, binomial expansion, completing the square or first-order differential equation analysis manipulation, inter alia, prior to classification of the function into one of the four models for use in the method of the present invention.

In light of the above, it is therefore understood that the invention may be practiced otherwise than as specifically described.

Attorney Docket No. 79236

ENHANCED MODEL IDENTIFICATION IN SIGNAL PROCESSING USING ARBITRARY EXPONENTIAL FUNCTIONS

ABSTRACT OF THE DISCLOSURE

A method for finding a probability density function (PDF) and its statistical moments for a chosen one of four newly derived probability models for an arbitrary exponential function of the forms $g(x) = \alpha x^m e^{-\beta x^n}$, $-\infty < x < \infty$; $g(x) = \alpha x^m e^{-\beta x^n}$, $0 \le x < \infty$; $g(x) = \alpha\left(\frac{x-a}{b}\right)^m e^{-\beta\left(\frac{x-a}{b}\right)^n}$, $-\infty < x < \infty$; and $g(x) = \alpha\left(\frac{x-a}{b}\right)^m e^{-\beta\left(\frac{x-a}{b}\right)^n}$, $0 \le x < \infty$. The model chosen will depend on the domain of the data and whether information on the parameters *a* and *b* exists. These parameters may typically be the mean or average of the data and the standard deviation, respectively. Non-linear regression analyses are performed on the data distribution and a basis function is reconstructed from the estimates in the final solution set to obtain a PDF, a moment generating function and the mean and variance. Simple hypotheses about the behavior of such functional forms may be tested statistically once the empirical least squares methods have identified an applicable model derived from actual measurements.

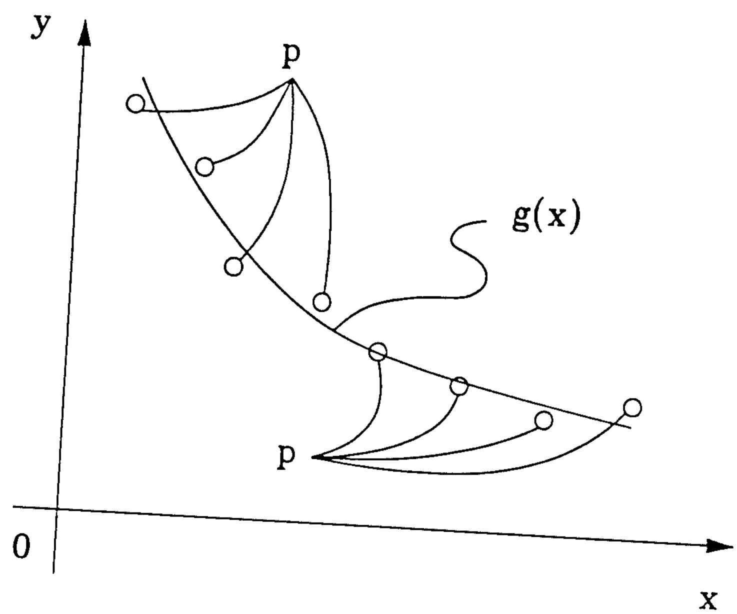

FIG. 1

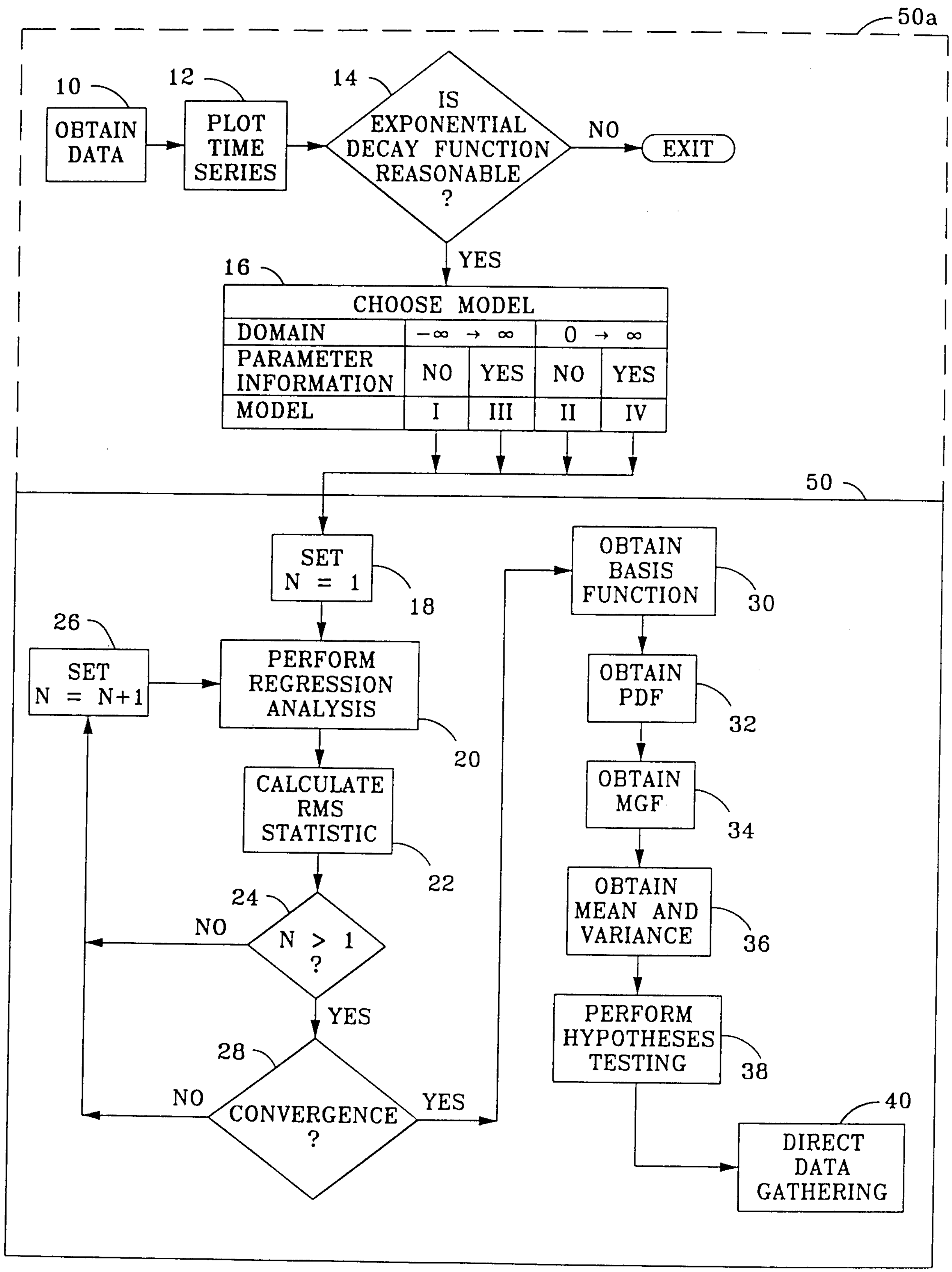
50a
10
OBTAIN DATA
12
PLOT TIME SERIES
14
IS EXPONENTIAL DECAY FUNCTION REASONABLE ?
NO
EXIT
YES
16
CHOOSE MODEL
DOMAIN
−∞ → ∞
0 → ∞
PARAMETER INFORMATION
NO
YES
NO
YES
MODEL
I
III
II
IV
50
SET N = 1
18
26
SET N = N+1
PERFORM REGRESSION ANALYSIS
20
CALCULATE RMS STATISTIC
22
24
NO
N > 1 ?
YES
28
NO
CONVERGENCE ?
YES
OBTAIN BASIS FUNCTION
30
OBTAIN PDF
32
OBTAIN MGF
34
OBTAIN MEAN AND VARIANCE
36
PERFORM HYPOTHESES TESTING
38
40
DIRECT DATA GATHERING

FIG. 2