Differential ant-stigmergy algorithm based parameters estimation of sum of exponentials model

Jianfeng Liu , Jiawen Bian, Hongwei Li

***Abstract*—**Exponentials sum model is an important model in time series analysis and has applications in modeling various physical phenomena of real life. The estimation of the parameters of the model is a necessary and fundamental task for the application of the model. In this paper, we propose a differential ant-stigmergy algorithm (DASA) based iterative procedure to estimate the parameters of the considered model where two different criterions which are least squares and least absolute errors are considered. The estimators of the parameters for the considered model by the two criterions are compared with the existing estimators by genetic algorithm based least squares (GA-LS). Simulation experiments and real data fitting are presented to inspect the performance of the proposed algorithm. It can be observed that better results can be obtained by DASA based LS (DASA-LS) than by GA-LS in terms of mean squares errors (MSEs) and robustness. Although a higher dimensional optimization is needed for DASA based least absolute errors (DASA-LAE) than that for the other two methods, DASA-LAE providers better results than DASA-LS in outliers condition. Finally, simulation results also show that DASA has better global searching ability than that for GA.

***Keywords*—**Differential ant-stigmergy algorithm, Least absolute errors estimation, Least square estimation , Sum of exponentials model.

I. Introduction

In this paper, we study the problem of the parameters estimation of the sum of the exponentials model as follows:

$$y(t)=\sum_{k=1}^{p}\alpha_k e^{\beta_k t}+\varepsilon(t) \tag{1}$$

This work is supported partially by the National High Technology Research and Development Program ("863"Program) of China (Grant number: 2014AA06A608), the National Nature Science Foundation of China (Grant number: 61302138), the Open Project from Hubei Key Laboratory of Intelligent Geo-information processing (Nos. KLIGIP2016A01, KLIGIP2016A02) and the Nature Science Foundation of Hubei Province (Grant number: 2015CFB555).

Jianfeng Liu, Jiawen Bian, and Hongwei Li are all with the School of Mathematics and Physics, China University of Geosciences, Wuhan, 430074, China (phone:86-27-67883092); email: liujianf206@163.com, jwbian@cug.edu.cn, hwli@cug.edu.cn

where $\{y(t)\}$ are observed data at time point $t\in\{t_1,t_2,\cdots,t_n\}$ which can be equidistant or non-equidistant. $\{\alpha_k\}_{k=1}^{p}$ and $\{\beta_k\}_{k=1}^{p}$ are unknown amplitude and frequency parameters and p is the number of components in model (1) and is assumed to be known. The sequence of additive noise $\{\varepsilon(t)\}$ is assumed to be a sequence of independent and identically distributed (i.i.d.) random variable with zero mean and finite variance σ^2. The focus of the paper is to estimate the parameters $\{\alpha_k\}_{k=1}^{p}$ and $\{\beta_k\}_{k=1}^{p}$ according to the given observations $\{y(t_i)\}_{i=1}^{n}$ where n is the size of the sample.

Exponentials sum model has important and extensive applications in many practical problems such as compartmental modeling and tracer kinetics [1], transmission function analysis in radiative transmission research [2], gene expressions [3], fitting of dwell-time distributions [4] and modeling probability densities [5]. Other literatures related to this model are referred to [6]-[8] and the references therein.

It is a fundamental problem for estimating the parameters of sum of exponentials model. There are many different ways of estimating parameters of sum of exponentials model. Among non-parametric methods, Prony algorithm is the first attempt to deal with the estimation of the parameters of exponentials sum model [9]. In Prony method, the solution of the exponentials sum model was interpreted as a solution to the homogeneous ordinary differential equation. The solution precision is reduced largely when the signal is imbedded in noise. The modification of Prony algorithm is referred to [10]-[11].

It is known that least squares (LS) criterion can produce good estimators and have good statistical properties for this problem in white noise condition [12], however, the methods above are not directly to deal with the objective function of the LS. Moreover, it is also known to be numerically difficult to solve the LS estimators [12].

In addition to LS, LAE is another widely used criteria in optimization such as compressed sensing [13], identification of isolated structural damage [14] and impedance inversion [15]. LAE is also known to produce better estimation than LS in the

outliers condition, however, the objective function of LAE is the absolute value function which brings the trouble for Newton based optimization algorithms and can not use the separation regression technique to reduce the optimization dimension. It is noted that multidimensional optimization is very complicated problem [16], which makes the optimization algorithms to search in a wider and more complex space. A lot of excellent optimization algorithm show bad performance as increasing of the variable dimension [16]. So LAE seems to be a numerically more difficult problem than LS. To the best of our knowledge, there are not relevant literatures discussing the optimization of exponentials sum model based on LAE.

It is known that traditional optimization algorithms such as [17] applying standard least squares procedures often ended up with derivative and matrix related computational problems such as premature termination to the search for the minimum, unreliable covariance matrices for the estimates and failure to give adequate error indications. Several special purpose algorithms have been proposed [18]-[23] to overcome these problems.

Bionic optimization algorithm is a kind of excellent optimization algorithm which has very strong versatility and is not bound by specific practical problems and avoids many of the shortcomings mentioned above [24]-[31]. Genetic algorithm is one of intelligent optimization algorithms which was introduced to solve the parameter estimation of model (1) by [32]. This algorithm directly deals with the objective function by LS. However, large number of individuals in group, gene expression and transformation between the coding and decoding require large number of calculation cost.

Korošec designed the differential ant-stigmergy algorithm (DASA) which can efficiently solve high-dimensional, wide search range, nonlinear optimization problems [33] and hasn't been applied to model (1). Inspired by this algorithm, in this paper, we combine DASA with LS as well as LAE to estimate the parameters of model (1).

The paper is arranged as follows. First, the optimization method of DASA is introduced and the algorithm of DASA-LS and DASA-LAE are described in Section Ⅱ. In Section Ⅲ, simulation experiments are performed to verify the effectiveness of the proposed algorithm. An application of the proposed algorithm on real measured data is presented in Section Ⅳ. Finally, we conclude the paper in Section Ⅴ.

II. Proposed Algorithm

A. DASA

The DASA is an ant-colony optimization based algorithm for numerical optimization capable of solving high-dimensional real-parameter optimization problems. In this section, we will use DASA-LS and DASA-LAE to estimate the parameters of the considered model (1).

For a given objective function $f(\boldsymbol{x})$ which is to be optimized with respect to $\boldsymbol{x}=(x_1,x_2,\cdots,x_D)^{\mathrm{T}}$, the aim of DASA is to find a vector of parameter values $\boldsymbol{x}^*=(x_1,x_2,\cdots,x_D)^{\mathrm{T}}$ that minimizes $f(\boldsymbol{x})$, i.e.,

$$\boldsymbol{x}^*=\arg\min_{x\in R^D} f(\boldsymbol{x}) \tag{2}$$

Because DASA uses the idea of ant colony algorithm, the establishment of a graph and distribution, update of pheromone and selection of feasible solution elements are included. Next we will introduce concepts of difference, differential graph, and assignment of pheromones for the ultimate DASA.

To solve this real-parameter optimization problem, a fine-grained discrete form of continuous domain is created. Using this form, we are able to represent the problem as a graph. In order to set up a graph, parameter difference is introduced in DASA which can become a vertex in the graph. Let x_i be the current value of the *i*-th parameter. According to its value range and the preset maximum precision of the parameter, a series of parameter differences can be produced and some of them can be used to update x_i in each iteration:

$$x_i'=x_i+\delta_i \tag{3}$$

δ_i is a parameter difference and x_i' is the updated value of x_i. δ_i has $2d_i+1$ possible values, where $d_i=U_i-L_i+1$, $U_i=\lfloor \log_b(\max(x_i)-\min(x_i))\rfloor$, $L_i=\lfloor \log_b(\varepsilon_i)\rfloor$ and $\lfloor\bullet\rfloor$ is rounding down function. With the parameter ε_i, the maximum precision of parameter x_i is set and b is the discrete base.

Once we get parameter difference, a differential graph can be established as follows by considering all possible values of $\{\delta_i\}_{i=1}^{D}$ (see Fig. 1).

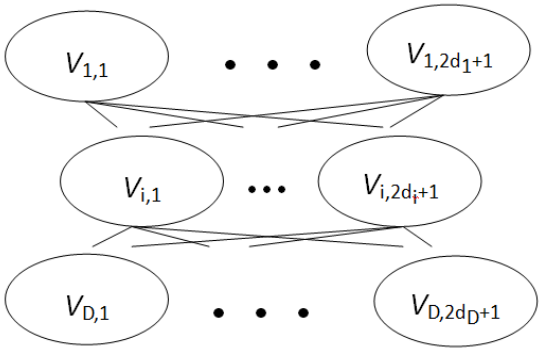

Fig. 1 A schematic representation of a differential graph

In Fig. 1, $\{V_{i,j}\}_{j=1}^{2d_i+1}$ denote the vertices of the graph relevant to the *i*-th parameter x_i. The value for $V_{i,j}$ is *j*-th value of δ_i. This graph is divided into *D* layers. Each vertex of upper layer is connected to all the vertices that belong to the next layer. DASA introduces *m* artificial ants and each ant randomly selects a vertex in each layer in turn from the first layer to the *D*-th layer. Let $\boldsymbol{V}=(V_1,V_2,\cdots,V_D)^{\mathrm{T}}$ be one path that an ant select,

parameter vector $\boldsymbol{x}=(x_1,x_2,\cdots,x_D)^{\mathrm{T}}$ can be updated according to $\boldsymbol{V}$ and (3).

In addition to the above two operations, the assignment of pheromones on the vertices is indispensable. The deposition of pheromones to the vertices of the differential graph is made according to Cauchy distribution as follow

$$C(z)=\left[s\pi\left(1+\left(\frac{z-l}{s}\right)^2\right)\right]^{-1} \tag{4}$$

where l is the location offset and $s=s_{global}-s_{local}$ is the scale factor. For the initial pheromone distribution, let $l=0, s_{global}=1, s_{local}=0$ and each layer of parameter vertices is equidistantly arranged between $z=$[-4,4]. The minimum and maximum *z* values were chosen according to the behavior of the standard Cauchy distribution, where at -4 and 4, the probability is already close to zero, so choosing higher values would be useless.

The implementation of DASA can be described as follows:

Step 1. Set the number of ant m, the pheromone evaporation factor ρ, the global scale-increasing factor s_+, the global scale-decreasing factor s_-, s_{global}, s_{local}, discrete base b and maximum precision ε_i ($i=1,2,\cdots,D$). Construct the differential graph and deposit pheromones on all vertices according to (4).

Step 2. Generate temporary solution $\boldsymbol{x}^{tbest}$ and calculate $y^{tbest}=f(\boldsymbol{x}^{tbest})$. Simultaneously, Let $\boldsymbol{x}^{best}=\boldsymbol{x}^{tbest}$ and $y^{best}=f(\boldsymbol{x}^{best})$.

Step 3. Each ant begins to select path from the starting vertex which is in first layer with a corresponding probability

$$\mathrm{prob}(V_{i,j})=\tau(V_{i,j})\Big/\sum_{k=1}^{2d_i+1}\tau(V_{k,j})$$

where $\tau(V_{i,j})$ is the pheromone amount on the vertex $V_{i,j}$.

Step 4. Get m solutions based on the same $\boldsymbol{x}^{tbest}$ and m paths according to (3). Calculate the corresponding m values of objective function. Let $y^{cbest}=f(\boldsymbol{x}^{cbest})$ denotes the minimal value among all *m* values where $\boldsymbol{x}^{cbest}=\boldsymbol{x}^{tbest}+\boldsymbol{p}^{cbest}$ and $\boldsymbol{p}^{cbest}$ is the corresponding path from $\boldsymbol{x}^{tbest}$ to $\boldsymbol{x}^{cbest}$. If $\boldsymbol{x}^{cbest}$ is better than $\boldsymbol{x}^{tbest}$, then $\boldsymbol{x}^{tbest}=\boldsymbol{x}^{cbest}$. At the same time, $s_{global}=(1+s_+)s_{global}$, $s_{local}=\frac{1}{2}s_{global}$ and the pheromone amount is redistributed according to associated path $\boldsymbol{p}^{cbest}$ and renewing $s=s_{global}-s_{local}$. Furthermore, if the new $\boldsymbol{x}^{tbest}$ is better than the global best solution $\boldsymbol{x}^{best}$, then $\boldsymbol{x}^{best}=\boldsymbol{x}^{tbest}$ and renewing y^{best} accordingly. If *m* ants do not generate the better solution than $\boldsymbol{x}^{tbest}$, then $s_{global}=(1-s_-)s_{global}$, $l=(1-\rho)l$ and $s_{local}=(1-\rho)s_{local}$, so the pheromone amount on each vertex can be updated.

Step 5. If the pre-set number of iterations is reached, the optimization procedure is terminated. Otherwise, the procedure is returned to step 3.

Remark1: In step 1, $l=0$, $s_{global}=1$, $s_{local}=0$ and they are changed along with the iteration. ρ, b, ε_i are kept constant throughout the iteration.

Remark2: As we can see that the pheromone on each layer must be updated in Fig. 1. When $\boldsymbol{p}^{cbest}$ in step 4 is obtained, we find the location in [-4,4] of the first component of $\boldsymbol{p}^{cbest}$ which is the value of l in (4) needed to update the first layer of pheromone. According to the same method, l in (4) needed to update the rest layers of pheromone can be obtained.

Remark3: When (4) is used to arrange pheromone on differential graph, the value of l is different on different layer while the value of s is the same for all layers.

B. DASA-LS estimation of the parameters of sum of exponentials model

In this section, we will use DASA-LS to estimate the parameters of the sum of exponentials model.

It is known that once the parameters $\{\beta_k\}_{k=1}^p$ are estimated, model (1) can be changed to a linear model of $\{\alpha_k\}_{k=1}^p$ which can be obtained by the linear regression technique [32].

Observe that under the assumption of additive noise, the LS estimators can be obtained by minimizing the residual sums of squares with respect to $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$,where

$\boldsymbol{\alpha}=(\alpha_1,\alpha_2,\cdots,\alpha_p)^{\mathrm{T}}$ and $\boldsymbol{\beta}=(\beta_1,\beta_2,\cdots,\beta_p)^{\mathrm{T}}$

$$R(\boldsymbol{\alpha},\boldsymbol{\beta})=\sum_{t=t_1}^{t_n}\left[y(t)-\sum_{k=1}^{p}\alpha_k e^{\beta_k t}\right]^2 \tag{5}$$

For simplicity, (5) can be written in matrix notation as follows:

$$R(\boldsymbol{\alpha},\boldsymbol{\beta})=\left(\boldsymbol{Y}-\boldsymbol{A}(\boldsymbol{\beta})\boldsymbol{\alpha}\right)^{\mathrm{T}}\left(\boldsymbol{Y}-\boldsymbol{A}(\boldsymbol{\beta})\boldsymbol{\alpha}\right) \tag{6}$$

where $\boldsymbol{Y}=(y(t_1),y(t_2),\cdots,y(t_n))^{\mathrm{T}}$ and $\boldsymbol{A}(\boldsymbol{\beta})$ is $n\times p$ matrix of the following form:

$$\boldsymbol{A}(\boldsymbol{\beta})=\begin{pmatrix} e^{\beta_1 t_1} & \cdots & e^{\beta_p t_1} \\ \vdots & & \vdots \\ e^{\beta_1 t_n} & \cdots & e^{\beta_p t_n} \end{pmatrix} \tag{7}$$

Observe that for a fixed $\boldsymbol{\beta}$, the LS estimator of $\boldsymbol{\alpha}$ is

$$\tilde{\boldsymbol{\alpha}}(\boldsymbol{\beta})=\left[\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{\beta})\boldsymbol{A}(\boldsymbol{\beta})\right]^{-1}\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{\beta})\boldsymbol{Y} \tag{8}$$

Replacing $\boldsymbol{\alpha}$ in (6) with $\tilde{\boldsymbol{\alpha}}(\boldsymbol{\beta})$, we obtain

$$R(\tilde{\boldsymbol{\alpha}},\boldsymbol{\beta})=Q(\boldsymbol{\beta})=\boldsymbol{Y}^{\mathrm{T}}\left(\boldsymbol{I}-\boldsymbol{P}_{\boldsymbol{A}}\right)\boldsymbol{Y} \tag{9}$$

where $\boldsymbol{P}_{\boldsymbol{A}} = \boldsymbol{A}(\boldsymbol{\beta})\left[\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{\beta})\boldsymbol{A}(\boldsymbol{\beta})\right]^{-1}\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{\beta})$ is the projection matrix on the space spanned by the columns of $\boldsymbol{A}(\boldsymbol{\beta})$. So the LS estimator of $(\boldsymbol{\alpha},\boldsymbol{\beta})$ obtained by minimizing $R(\boldsymbol{\alpha},\boldsymbol{\beta})$ with respect to $\boldsymbol{\alpha},\boldsymbol{\beta}$ is the same as first obtaining the LS estimator of $\boldsymbol{\beta}$ by minimizing $Q(\boldsymbol{\beta})$ with respect to $\boldsymbol{\beta}$ and then obtaining the LS estimator of $\boldsymbol{\alpha}$ according to the estimated $\boldsymbol{\beta}$ in (8).

Incorporating DASA in the proposed algorithm, the parameter vector to be optimized is $\boldsymbol{\beta}$, i.e. $\boldsymbol{x} = \boldsymbol{\beta}$, $D = p$. The objective function f is $Q(\boldsymbol{\beta})$ in (9). Once the estimator of $\boldsymbol{\beta}$ is obtained, estimator of $\boldsymbol{\alpha}$ can be obtained by (8).

C. DASA-LAE estimation of the parameters of sum of exponentials model

LAE is another rule of optimization which can be obtained by minimizing the absolute value distance between the observed data and model data with respect to $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ as follows:

$$L(\boldsymbol{\alpha},\boldsymbol{\beta}) = \sum_{t=t_1}^{t_n}\left|y(t) - \sum_{k=1}^{p}\alpha_k e^{\beta_k t}\right| \tag{10}$$

Since the separation regression technique can not be used due to the absolute value used in (10), frequency and amplitude are needs to be estimated simultaneously. So LAE estimation is a higher dimensional optimization problem than LS estimation in (9). However, LAE produces smaller error than LS for the same estimates of $\boldsymbol{\alpha},\boldsymbol{\beta}$, it seems to obtain better results for LAE than LS in the outliers condition. It is noted that Newton based optimization algorithms can not effectively solve this kind of problem as it is based on gradient and is easy to fall into local minima while operation of taking absolute value has nearly nothing effect for DASA. So we use DASA based LAE to obtain estimators of the frequencies and amplitudes of the considered model in this paper.

As in section 2.2, DASA-LAE estimation can be obtained according to the similar steps as for DASA-LS. It is noted again that frequency and amplitude parameters are simultaneously estimated, which cause the algorithm to be a higher dimensional optimization than that in DASA-LS. So the parameter vector to be optimized is $\left(\boldsymbol{\alpha}^{\mathrm{T}},\boldsymbol{\beta}^{\mathrm{T}}\right)^{\mathrm{T}}$ for DASA-LAE, i.e. $\boldsymbol{x} = \left(\boldsymbol{\alpha}^{\mathrm{T}},\boldsymbol{\beta}^{\mathrm{T}}\right)^{\mathrm{T}}$, $D = 2p$. The objective function f is $L(\boldsymbol{\alpha},\boldsymbol{\beta})$ in (10).

III. SIMULATION STUDIES

In this section, we will present simulation results to assess the performance of the proposed algorithm, i.e. DASA-LS and DASA-LAE. Different levels of noise, different sizes of sample and outliers are considered in the experiments.

The computer platform used to perform the experiments is based on Intel Core 2.67-GHz processor, 3.25 GB of RAM and the Microsoft Windows 7 Professional Operating System. All algorithms are implemented by Matlab R2009a.

A. Performance of DASA-LS for real compartment growth model

To investigate the performance of DASA-LS for various noise levels and sample sizes, we consider the following growth model:

$$y(t) = \alpha_1 e^{\beta_1 t} + \alpha_2 e^{\beta_2 t} + \varepsilon(t) \tag{11}$$

where $t = 1,\cdots,n$ and n is the size of the sample. The real parameter values in model (11) are taken as $\alpha_1 = -6$, $\alpha_2 = 3$, $\beta_1 = -0.2320$, and $\beta_2 = 0.0119$. $\varepsilon(t)$ is taken as independently and identically distributed (i.i.d) normal noise sequence. In order to consider the influence of the noise for the proposed algorithm, we consider different standard deviation $\sigma's$ which are $\sigma = 0.01$, 0.05, and 0.1 respectively.

The procedure for finding the estimates of the frequency parameter vector $\boldsymbol{\beta}$ and amplitude parameter vector $\boldsymbol{\alpha}$ is replicated 200 times. The experimental parameter settings for DASA-LS are presented in Table 1. We report the average estimates (AEs), the corresponding mean square errors (MSEs) and the average absolute percentage errors (AAPEs) in Table 2 corresponding to different noise situations where different sample sizes are also considered which are 25, 50, 75, and 100. The results inside the parentheses and brackets in Table 2 are MSEs and AAPEs respectively. From Table 2, it can be observed that the performance of DASA-LS is quite satisfactory about AEs, MSEs, and AAPEs. General trends of the results of the estimates become better as the noise variance becomes smaller for given sample size. More specifically, the estimates of AEs approach the real values while MSEs and AAPEs becomes smaller with the decrease of noise variance for given sample sizes.

Table 1. Parameter settings for DASA-LS and model (11)

DASA parameters	Values
Number of ants	2
Discrete bases	2
Maximum precision	2-15
Global scale-increasing factor	0.02
Global scale-decreasing factor	0.01
Evaporation factor	0.2
Range of parameters	$\beta_1 \in [-10,10], \beta_2 \in [-10,10]$
Maximum number of iteration	800

The results of DASA-LS are also compared with the GA-LS algorithm in [32]. The comparison between DASA-LS and GA-LS for β_1, β_2, α_1, and α_2 are presented in Fig. 2 where $\sigma = 0.1$. From Fig. 2, it can be observe that the MSEs of DASA-LS are smaller than those of GA-LS. Meanwhile the results in Table 2 and Fig. 2 are obtained under the condition when the hunting zone of the parameters is [-10 10] for DASA and [-2 2] for **GA**, which indicates DASA has good global searching ability in a wider range than that for GA. Besides this,

the number of individuals in the population is different where 250 individuals are needed in one generation in GA-LS while only 2 ants are needed in one generation in DASA-LS, which indicates the number of individual needed in one generation for DASA is less than that for GA.

In order to further demonstrate the performance of DASA-LS, the condition of $n=75$ and $\sigma=0.05$ is investigated. Here the range of parameters β are set in [-2 2]. The convergence process of objective function and the estimates of the parameters β are reported in Fig. 3 and Fig. 4 respectively. From Fig. 3 and Fig. 4, it is obvious that the AEs of the parameters approach the stable values which are very close to the corresponding true values as the value of the objective function continuously decreases and achieves stability.

We also study the effect of population size with respect to stability of the algorithm and present the results in Fig. 5. It is observed from Fig. 5 that all of the objective function values go to stabilize as the iteration number increases for different population sizes where the iteration numbers are all less than 200. Moreover the iteration number for stability of objective function value is different for different population sizes and the iteration number seems to decrease with the increase of the sample size.

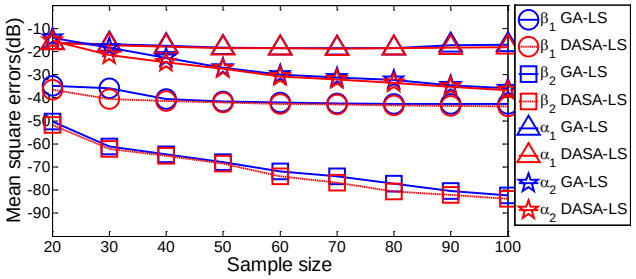

Fig. 2. Comparison of DASA-LS with GA-LS

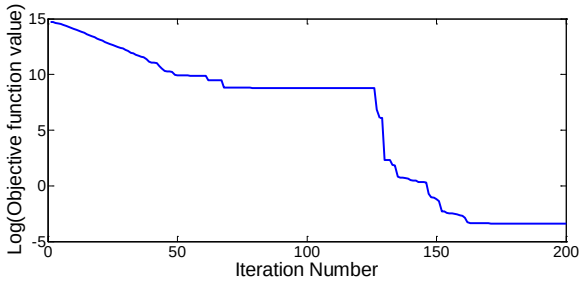

Fig. 3. Convergence graph of objective function for model (11)

Table 2. The performance of DASA-LS for parameter estimates of α and β in model (11)

σ	n	$\beta_1=-0.2320$	$\alpha_1=-6$	$\beta_2=0.0119$	$\alpha_1=3$
0.01	25	-0.2319 (2.38e-6) [0.52]	-6.0025 (2.71e-4) [0.22]	0.0119 (7.37e-8) [1.81]	3.0022 (3.18e-4) [0.47]
	50	-0.2320 (7.61e-7) [0.30]	-5.9999 (1.45e-4) [0.15]	0.0119 (1.56e-9) [0.26]	3.0000 (1.96e-5) [0.11]
	75	-0.2319 (5.28e-7) [0.24]	-5.9996 (1.70e-4) [0.17]	0.0119 (2.25e-10) [0.10]	3.0001 (6.40e-6) [0.06]
	100	-0.2319 (5.16e-7) [0.24]	-5.9988 (1.44e-4) [0.16]	0.0119 (8.79e-11) [0.06]	3.0000 (4.61e-6) [0.05]
0.05	25	-0.2320 (5.04e-5) [2.46]	-6.0006 (7.01e-3) [1.10]	0.0119 (1.47e-6) [8.04]	3.0006 (6.20e-3) [2.07]
	50	-0.2322 (1.70e-5) [1.44]	-6.0043 (4.30e-3) [0.80]	0.0119 (4.10e-8) [1.32]	2.9988 (5.10e-4) [0.57]
	75	-0.2323 (1.26e-5) [1.22]	-6.0033 (3.50e-3) [0.79]	0.0119 (5.02e-9) [0.47]	3.0001 (1.40e-4) [0.32]
	100	-0.2319 (9.48e-6) [1.01]	-5.9979 (3.00e-3) [0.73]	0.0119 (1.54e-9) [0.27]	3.0004 (7.75e-5) [0.23]
0.1	25	-0.2319 (2.14e-4) [5.00]	-6.0151 (2.86e-2) [2.17]	0.0118 (6.57e-7) [17.00]	3.0130 (2.97e-2) [4.56]
	50	-0.2325 (6.37e-5) [2.81]	-6.0009 (1.40e-2) [1.54]	0.0119 (1.41e-7) [2.53]	3.0017 (1.80e-3) [1.11]
	75	-0.2321 (5.25e-5) [2.46]	-6.0026 (1.37e-2) [1.52]	0.0119 (1.99e-8) [0.95]	2.9999 (5.65e-4) [0.64]
	100	-0.2326 (4.20e-5) [2.42]	-5.9982 (1.57e-2) [1.68]	0.0119 (4.23e-9) [0.53]	3.0006 (2.16e-4) [0.39]

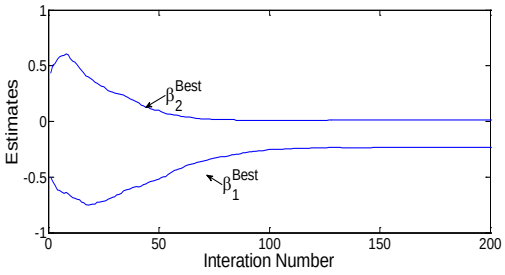

Fig. 4. Convergence graph of the estimates of parameter β_1 and β_2 for model (11)

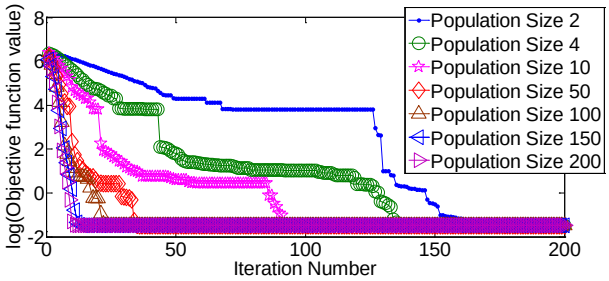

Fig. 5. Convergence graph of the objective function for different population sizes for model (11)

B. Performance of DASA-LS for real compartment growth model and DASA-LS for real compartment growth model with outliers

In order to further check the performance of DASA for outliers condition, we consider DASA-LS and DASA-LAE in model (11) where 10% observations which are modeled as outliers are randomly chosen to be added with noises from a Gaussian random variable having variance 4 while the other

90% observations are contaminated by Gaussian noise with variance σ^2 . Since DASA-LAE needs multidimensional optimization which simultaneously solves the estimates of frequency and amplitude, much time is needed for DASA-LAE than for DASA-LS . We set the maximum number of iteration to be 4000. The other experimental parameter settings for DASA-LS and DASA-LAE are the same as in Table 1. We report the AEs, MSEs, and AAPEs of DASA-LAE over 200 simulation runs in Table 3.

Table 3. The performance of DASA-LAE algorithm for parameter estimates of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ in outlier noise

σ	n	$\beta_1=-0.2320$	$\alpha_1=-6$	$\beta_2=0.0119$	$\alpha_1=3$
0.01	25	-0.2371 (8.71e-5) [3.35]	-5.9696 (6.41e-3) [0.90]	0.0129 (3.10e-6) [12.14]	2.9415 (1.22e-2) [3.04]
	50	-0.2329 (6.80e-6) [0.74]	-6.0138 (3.82e-3) [0.46]	0.0119 (8.43e-9) [0.62]	2.9965 (1.00e-4) [0.27]
	75	-0.2324 (4.32E-6) [0.53]	-6.0090 (4.51e-3) [0.41]	0.0119 (7.91e-10) [0.19]	2.9988 (2.31e-5) [0.13]
	100	-0.2324 (3.62E-6) [0.45]	-6.0083 (3.71e-3) [0.35]	0.0119 (2.27e-10) [9.82e-2]	2.9994 (1.13e-5) [8.71e-2]
0.05	25	-0.2452 (8.74e-4) [8.85]	-6.0730 (0.24) [3.33]	0.0136 (1.43e-5) [26.12]	2.9023 (5.61e-2) [6.57]
	50	-0.2344 (5.50e-5) [2.41]	-6.0239 (1.34e-2) [1.19]	0.0120 (1.14e-7) [2.26]	2.9907 (1.51e-3) [1.01]
	75	-0.2320 (3.00e-5) [1.85]	-5.9927 (1.30e-2) [1.35]	0.0119 (1.11e-8) [0.70]	2.9986 (3.33e-4) [0.49]
	100	-0.2330 (8.61e-5) [1.89]	-6.0131 (9.42e-2) [1.79]	0.0119 (3.10e-9) [0.36]	2.9976 (1.59e-4) [0.33]
0.1	25	-0.2343 (8.91e-4) [10.30]	-6.0921 (0.12) [4.09]	0.0118 (3.12e-5) [37.00]	3.0416 (1.63e-1) [10.02]
	50	-0.2343 (2.32e-4) [4.72]	-6.0347 (0.10) [2.62]	0.0119 (3.82e-7) [4.30]	2.9965 (5.01e-3) [1.91]
	75	-0.2357 (1.43e-4) [3.91]	-6.0893 (0.11) [3.03]	0.0119 (5.27e-8) [1.51]	3.0099 (1.41e-3) [0.97]
	100	-0.2343 (1.33e-4) [3.80]	-6.0578 (0.11) [2.79]	0.0119 (1.31e-8) [2.88]	3.9944 (6.50e-4) [0.69]

From Table 3, it can be observed that although the noise variance of outlier is much larger than the noise variance of other observations where the ratio of variance between outlier and other observations are 400,1600, and 40000 for the three noise conditions considered, the performance of the estimators are still very satisfactory. MSEs and AAPEs become small with the increase of sample sizes while decrease when the noise standard deviations decrease. Moreover, comparing Table 3 with Table 2， it is observed that the estimates in each table are very close with each other , which indicates DASA-LAE is robust for outliers.

In addition, we plot the MSEs of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ for DASA-LS and DASA-LAE where n =75, σ =0.01 over 200 simulation runs in Fig. 6. From Fig. 6, it can be observed that the red lines which denote the estimates of DASA-LSE are below the corresponding blue lines which denote the estimates of DASA-LS. So the performances of DASA-LAE are superior to the results based on DASA-LS.

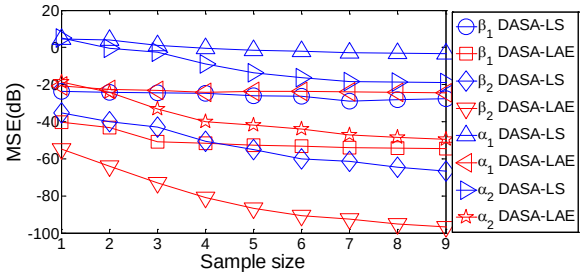

Fig. 6. Comparison of DASA-LAE and DASA-LS for β_1, β_2, α_1 and α_2

C. Performance of GA-LS, DASA-LS, and DASA-LAE for real compartment decay model

In this section, to further compare the performance of GA-LS, DASA-LS, and DASA-LAE, we consider a real compartment decay model which is another exponentials sum model as follows:

$$y(t_i)=\alpha_0+\alpha_1 e^{\beta_1 t_i}+\alpha_2 e^{\beta_2 t_i}+\varepsilon(t_i) \quad (12)$$

where $t_i=i/n, i=1,\cdots,n$ and n is the size of the sample. The true parameter values in model (12) are as follows: $\alpha_0=0.5$, $\alpha_1=-1.5$, $\alpha_2=2$, $\beta_1=-7$, and $\beta_2=-4$. The additive noise $\{\varepsilon(t_i)\}$ is a sequence of i.i.d Gaussian random variables, i.e. $\varepsilon(t_i)\sim N(0,\sigma^2)$ where σ^2 is the noise variance.

We estimate β_1, β_2, α_0, α_1, and α_2 by DASA-LS and DASA-LAE respectively. The experimental parameter settings are the same as in Table 1 except the maximum number of iteration being taken as 4000 for DASA-LS and 40000 for DASA-LAE. Since 5-dimention optimization is needed for DASA-LAE about $(\alpha_0,\alpha_1,\alpha_2,\beta_1,\beta_2)$ simultaneously, the maximum number of iteration is taken larger than that in DASA-LS.

In order to check the performance of DASA, we consider $\sigma=0.001$ and take different sample sizes as $n=50$, 100, 150, 200, 250, 300, 350, and 400. We report the MSEs of (β_1,β_2) of DASA-LS, DASA-LAE, and GA-LS over 200 independent runs in Fig. 7.

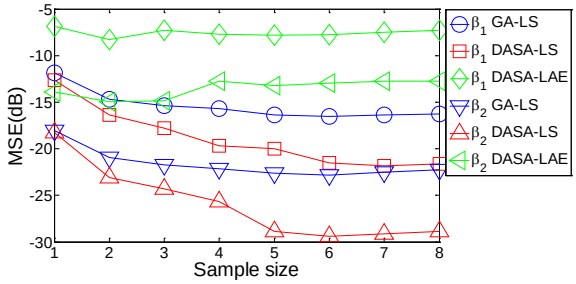

Fig. 7. Comparison of DASA-LAE, DASA-LS, and GA-LS

From Fig. 7, it is observed that the estimates of DASA-LS are better than those for GA-LS and DASA-LAE while the

estimates of DASA-LAE are worse than those for the other two estimates of GA-LS and DASA-LS.

In order to further study the performance of DASA-LAE, we consider the case when $n=100$ and $\sigma=0.001$. The AEs of β_1 and β_2 of DASA-LAE over 200 simulation runs are reported in Fig. 8. From Fig. 8, it is observed that although the MSEs of DASA-LAE are larger than those of another two methods in Fig. 7, the AEs of β_1 and β_2 are satisfactory and are very close to the true parameter values.

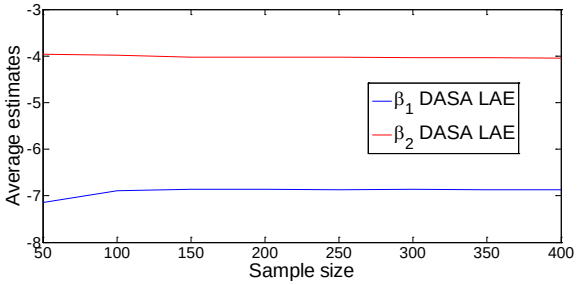

Fig. 8. The average estimates of β_1 and β_2 of DASA-LAE

IV. REAL LIFE DATA ANALYSIS

To check the performance of DASA-LS and DASA-LAE for real life data, two sets of real life data are fitted using sum of exponentials model.

The experimental parameters settings for the two real life analysis are presented in Table 4.

Table 4. Experimental parameter settings for real life data in [32]

DASA parameters	Values
The number of ant	2
Discrete base	2
Maximum precision	2^{-15}
Global scale-increasing factor	0.02
Global scale-decreasing factor	0.01
Evaporation factor	0.2
Maximum number of iteration	5000

Table 5. Sulfisoxazole concentration versus time [32]

Time(min)	Sul(*ug/ml*)	Time (min)	Sul (*ug/ml*)
0.25	215.6	3.00	101.2
0.5	189.2	4.00	88.0
0.75	176.0	6.00	61.6.00
1.00	162.8	12.00	22.00
1.50	138.6	24.00	4.4
2.00	121.0	48.00	0.1

The first set of data was obtained in [34] as presented in Table 5. It is on the metabolism of sulfisoxazole which was used in [8].

For this intravenous data, a 2-compartment model was proposed in [32]. The concrete form is:

$$C_1(t)=\alpha_1 e^{\beta_1 t}+\alpha_2 e^{\beta_2 t} \quad (13)$$

where $(\alpha_1,\alpha_2,\beta_1,\beta_2)$ is the parameter vector to be estimated and t is the time. Sulfisoxazole concentration at the time t is predicted by $C_1(t)$. It is noted that the sampling interval is not equidistant.

We take α_1,α_2 in [50 200] and β_1,β_2 in [-10 0] and use the DASA-LS and DASA-LAE to fit this set of data respectively. The fitted models through the DASA-LS and DASA-LAE are:

$$C_{1DASA-LS}(t)=81.242e^{-1.306t}+162.5971e^{-0.1618t} \quad (14)$$

$$C_{1DASA-LAE}(t)=84.5473e^{-1.2744t}+160.3874e^{-0.1597t} \quad (15)$$

The fitted model by the GA-LS in [32] is:

$$C_{1GA-LS}(t)=80.81e^{-1.31t}+162.93e^{-0.16t} \quad (16)$$

The sum of square of errors between fitting data and real measured data of DASA-LS, DASA-LAE, and GA-LS are 34.3765, 36.2282, and 37.2066 respectively. It is obvious that the result of DASA-LS is best. We also illustrate the relation between fitted data and real observed data in Fig. 9. The residual between fitted data and real observed data is graphically shown in Fig. 10. From Fig. 9, it is observed that the estimated models by the proposed algorithm can fit the observed data very well. From Fig. 10, the estimated residual of DASA-LS is smallest as a whole while the estimated residual of DASA-LAE is smaller than that of GA-LS.

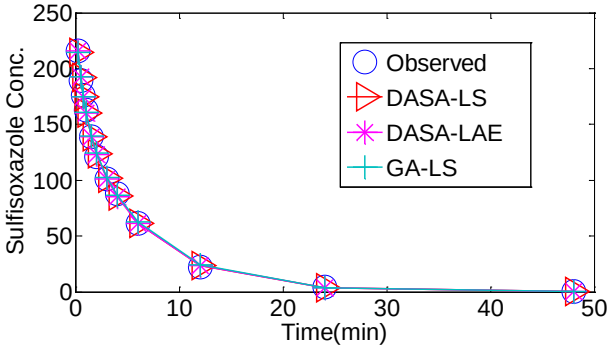

Fig. 9. Fit of the sulfisoxazole concentration data based on DASA-LS, DASA-LAE, and GA-LS.

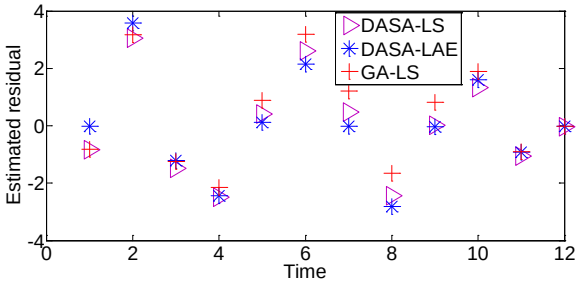

Fig. 10. Residual plot corresponding to DASA-LS, DASA-LAE, and GA-LS fit of the sulfisoxazole data

The second set of data is from [35] as reported in the following form (t, y_t):

(0; 0.844); (10; 0.908); (20; 0.932); (30; 0.936); (40; 0.925); (50; 0.908); (60; 0.881); (70; 0.850); (80; 0.818); (90; 0.784); (100; 0.751); (110; 0.718); (120; 0.685); (130; 0.658); (140; 0.628); (150; 0.603); (160; 0.580); (170; 0.558); (180; 0.538); (190; 0.522); (200; 0.506); (210; 0.490); (220; 0.478); (230; 0.467); (240; 0.457); (250; 0.448); (260; 0.438); (270; 0.431); (280; 0.424); (290; 0.420); (300; 0.414); (310; 0.411); (320; 0.406).

For this real life data, we use a 2-compartment model to fit the given data as in [32]:

$$C_2(t) = Ae^{\theta_1 t} + Be^{\theta_2 t} + C \quad (17)$$

where $(A, B, C, \theta_1, \theta_2)$ is the parameter vector to be estimated and t is the time. The real life data at time t is predicted by $C_2(t)$.

We take θ_1, θ_2 in $[-2, 0]$ and A, B, C in $[-4, 4]$. The fitted models by the DASA-LS and DASA-LAE are:

$$C_{2DASA-LS}(t) = 0.3754 - 1.4654e^{-0.02212t} + 1.9365e^{-0.01287t} \quad (18)$$

$$C_{2DASA-LAE}(t) = 0.3770 - 1.6361e^{-0.02169t} + 2.1032e^{-0.01318t} \quad (19)$$

The fitted model by the GA-LS in [32] is:

$$C_{2GA-LS}(t) = 0.3754 - 1.4615e^{-0.02213t} + 1.9327e^{-0.0129t} \quad (20)$$

The MSEs between fitting data and real measured data is 2.9865×10^{-9} by DASA-LS, however, this result is 8.51×10^{-9} by GA-LS. We also illustrate the relation between fitted data and real observed data in Fig. 11. The plot of residual between fitted data and real observed data is given in Fig. 12.

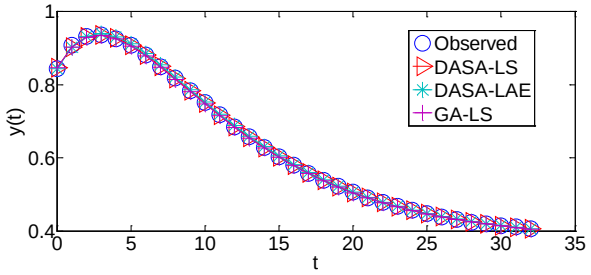

Fig. 11. Fit of the dataset of [35] based on DASA-LS,DASA-LAE, and GA-LS

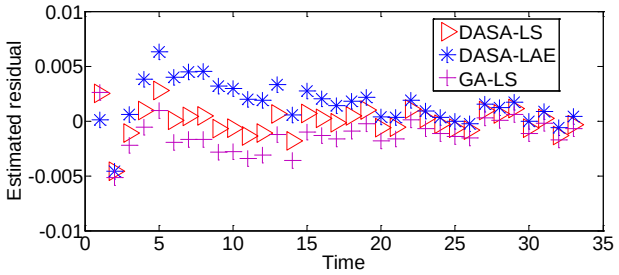

Fig. 12. Residual plot of DASA-LS, DASA-LAE, and GA-LS for the dataset of [35]

From Fig. 11 and Fig. 12, it can be observed that the fitting of DASA-LS is better than that of DASA-LAE and GA-LS. The fitting based on DASA-LAE is a little bit worse, however, its overall performance is satisfactory. The fitting data are very close to the true observations. The estimated residual based on DASA-LAE is of the same order of magnitude with that of the other two methods.

V. CONCLUSIONS

This paper proposed a DASA based procedure to estimate the parameters of sum of exponentials model. For the DASA based estimation, two estimation criterions which are least squares and least absolute errors are considered for white noise and outliers condition respectively. We compared the performance of estimators by DASA-LS and DASA-LAE with that of GA-LS through several simulation tests and real life data fitting. DASA is observed to be efficient in searching the optimal solutions with a very small number of ants and only two ants are used for the proposed algorithm. On the other hand, DASA can search the optimal solution in a larger range than GA while is robust for the initial value. Simulation studies show that DASA-LS based estimators have smaller MSEs than those by DASA-LAE and GA-LS in moderate additive noise condition. Moreover, although DASA-LAE needs a higher dimensional optimization than DASA-LA and GA-LS, it produces better results than DASA-LS in outliers condition which verified the effectiveness of DASA-LAE for outliers condition. Finally, analysis of real data show superiority of DASA-LS to GA-LS in terms of MSEs. The generated model based on DASA-LS can fit the measured data more accurate than GA-LS and DASA-LAE.

REFERENCES

[1] D. H. Anderson, "Compartmental Modeling and Tracer Kinetics," *Lecture Notes in Biomathematics*, Springer, New York, 1983.

[2] W. J. Wiscombe and J.W. Evans, "Exponential-sum fitting of radiative transmission functions," *Journal of Computational Physics*, vol. 24, no. 4, pp. 416-444, 1977.

[3] C. D. Giurcaneanu, I. Tabuş and J. Astola, "Clustering time series gene expression data based on sum-of-exponentials fitting," *Eurasip Journal on Advances in Signal Processing*, vol. 2005, no. 8, pp. 1159-1173, 2005.

[4] D. Landowne, B. Yuan and K. L. Magleby, "Exponential sum-fitting of dwell-time distributions without specifying starting parameters," *Biophysical Journal*, vol. 104, no. 11, pp. 2383-2391, 2013.

[5] B. Dumitrescu, B. C. Şicleru and F. Avram, "Modeling probability densities with sums of exponentials via polynomial approximation," *Journal of Computational and Applied Mathematics*, vol. 292, no. 15, pp. 513-525, 2016.

[6] A. Seber and B. Wild, Nonlinear Regression, Wiley, New York, 1989.

[7] A. R. Gallant, Nonlinear Statistical Models, Wiley, New York, 1987.

[8] D. M. Bates and D. G. Watts, Nonlinear Regression Analysis and its Applications, Wiley, New York, 1988.

[9] R. Prony Essai experimentale et analytique. J. Ecole Polytechnique Paris 1(2):24-76(1795)

[10] M. R. Osborne and G. K. Smyth, "A modified Prony algorithm for exponential function fitting," *SIAM Journal on Scientific Computing*, vol. 16, no. 1, pp. 119-138, 1995.

[11] D. Kundu and A. Mitra, " Estimating the parameters of linear compartment model," *Journal of Statistical Planning and Inference*, vol. 70, no. 2, pp. 317–334, 1998b.

[12] D. Kundu and A. Mitra, "Fitting a sum of exponentials to equispaced data," *Sankhya B*, vol. 60, no. 3, pp. 448-463, 1998a.

[13] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp.1289-1306, 2006.

[14] E. M. Hernandez, "Identification of isolated structural damage from incomplete spectrum information using L1 -norm minimization," *Mechanical Systems & Signal Processing*, vol. 46, no. 1, pp. 59-69,

2014.

[15] C. Liu, C. Song, Q. Lu, Y. Liu, X. Feng and Y. Gao, "Impedance inversion based on L1 norm regularization," *Journal of Applied Geophysics*, vol. 2015, no. 120, pp. 7-13, 2015.

[16] M. Lozano, D. Molina and F. Herrera, "Editorial scalability of evolutionary algorithms and other metaheuristics for large-scale continuous optimization problems," *Soft Computing*, vol. 15, no. 11, pp. 2085-2087, 2011.

[17] S. R. Wilson, "Benchmark data sets for the flexible evaluation of statistical software," *Computational Statistics and Data Analysis*, vol. 1, no. 1, pp. 29-39, 1983.

[18] R. H. Barham and W. Drane, "An algorithm for least squares estimation of non-linear parameters when some of the parameters are linear," *Technometrics*, vol. 14, no. 3, pp. 757-766, 1972.

[19] G. H. Golub and V. Pereyra, "The differentiation of pseudo-inverses and non-linear least squares problems whose variables separate," *SIAM Journal on Numerical Analysis*, vol. 10, no. 2, pp. 413-432, 1973

[20] M. R. Osborne, "Some special nonlinear least squares problems," *SIAM Journal on Numerical Analysis*, vol. 12, no. 4, pp. 571-592, 1975.

[21] M. R. Osborne, "Nonlinear least squares—the levenberg algorithm revisited," *Anziam Journal*, vol. 19. No. 3, pp. 343-357, 1976.

[22] M. R. Osborne and G. K. Smyth, "An algorithm for exponential fitting revisited," *Essays in Time Series and Allied Processes*, vol.1986, no. 23, pp. 419-430, 1986.

[23] M. Kahn, M. S. Mackisack, M. R. Osborne and G. K. Smyth, "On the consistency of Prony's method and related algorithms," *Journal of Computational and Graphical Statistics*, vol. 1, no. 4, pp. 329-349, 1992.

[24] Z. Cai, W. Gong, C. X. Ling, and H. Zhang, "A clustering-based differential evolution for global optimization," *Applied Soft Computing*, vol. 11, pp. 1363-1379, 2011.

[25] W. Gong, Z. Cai, C. X. Ling, and H. Li, "Enhanced Differential Evolution With Adaptive Strategies for Numerical Optimization," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 41, pp. 397-413, 2011.

[26] W. Gong and Z. Cai, "Differential evolution with ranking-based mutation operators," *IEEE transactions on cybernetics*, vol. 43, pp. 2066-2081, 2013.

[27] W. Gong, Z. Cai, and Y. Wang, "Repairing the crossover rate in adaptive differential evolution," *Applied Soft Computing*, vol. 15, pp. 149-168, 2014.

[28] W. Gong, A. Zhou, and Z. Cai, "A multioperator search strategy based on cheap surrogate models for evolutionary optimization," *IEEE Transactions on Evolutionary Computation*, vol. 19, pp. 746-758, 2015.

[29] W. Gong, Z. Cai, and D. Liang, "Adaptive ranking mutation operator based differential evolution for constrained optimization," *IEEE transactions on cybernetics*, vol. 45, pp. 716-727, 2015.

[30] C. Li, S. Yang, and M. Yang, "An adaptive multi-swarm optimizer for dynamic optimization problems," *Evolutionary computation*, vol. 22, pp. 559-594, 2014

[31] C. Li, S. Yang, and T. T. Nguyen, "A self-learning particle swarm optimizer for global optimization problems," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 42, pp. 627-646, 2012.

[32] S. Mitra, A. Mitra, "A genetic algorithms based technique for computing the nonlinear least squares estimates of the parameters of sum of exponentials model," *Expert Systems with Applications*, vol. 39, no. 7, pp. 6370-6379, 2012.

[33] P. Korošec, J. Šilc and B. Filipič, "The differential ant-stigmergy algorithm," *Information Sciences*, vol. 192, no. 5, pp. 82-97, 2012.

[34] S. A. Kaplan, R. E. Weinfeld, C. W. Abruzzo and M. Lewis, "Pharmacokinetic profile of sulfisoxazole following intravenous intramuscular and oral administration in man," *Journal of Pharmaceutical Sciences*, vol. 61, no. 5, pp. 773-778, 1972.

M. R. Osborne, "Some aspects of nonlinear least squares calculations," In: Lootsma FA (ed) Numerical methods for Nonlinear Optimization, Academic Press, New York, pp 171-189, 1972

Jianfeng Liu received his M.Sc. degree in applied mathematics in 2004 from Wuhan University. He is currently a lecturer at the School of Mathematics and Physics, China University of Geosciences, Wuhan. His research interests include statistical signal processing and time series analysis.

Jiawen Bian received his M.Sc. degree in applied mathematics in 2004 from Hubei University and Ph.D. degree in earth exploration and information technology in 2010 from the China University of Geosciences. He is currently an associate professor at the School of Mathematics and Physics, China University of Geosciences, Wuhan. His research interests include statistical signal processing and time series analysis.

Hongwei Li received the Ph.D. degree in applied mathematics from Peking University, Beijing, China, in 1996. From 1996 to 1998, he was a postdoctoral fellow at the Institute of Information Science, Beijing Jiaotong University. Since 1999, he has been a professor at the School of Mathematics and Physics, China University of Geosciences, Wuhan. His research interests include statistical signal processing, blind signal processing, multidimensional signal processing, pattern recognition and time series analysis.