

Research Article

A Method for Solving LiDAR Waveform Decomposition Parameters Based on a Variable Projection Algorithm

Ke Wang , Guolin Liu , Qiuxiang Tao, Luyao Wang, and Yang Chen

College of Geomatics, Shandong University of Science and Technology, Qingdao, Shandong 266590, China

Correspondence should be addressed to Ke Wang; wke0115@126.com and Guolin Liu; gliu@sdust.edu.cn

Received 12 January 2020; Accepted 4 May 2020; Published 20 May 2020

Academic Editor: Dimitri Volchenkov

Copyright © 2020 Ke Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Light detection and ranging (LiDAR) is commonly used to create high-resolution maps; however, the efficiency and convergence of parameter estimation are difficult. To address this issue, we evaluated the structural characteristics of received LiDAR signals by decomposing them into Gaussian functions and applied the variable projection algorithm of the separable nonlinear least-squares problem to the process of waveform fitting. First, using a variable projection algorithm, we separated the linear (amplitude) and nonlinear (center position and width) parameters in the Gaussian function model; the linear parameters are expressed with nonlinear parameters by the function. Thereafter, the optimal estimation of the characteristic parameters of the Gaussian function components was transformed into a least-squares problem only comprising nonlinear parameters. Finally, the Levenberg–Marquardt algorithm was used to solve these nonlinear parameters, whereas the linear parameters were calculated simultaneously in each iteration, and the estimation results satisfying the nonlinear least-square criterion were obtained. Five groups of waveform decomposition simulation data and ICESat/GLAS satellite LiDAR waveform data were used for the parameter estimation experiments. During the experiments, for the same accuracy, the separable nonlinear least-squares optimization method required fewer iterations and lesser calculation time than the traditional method of not separating parameters; the maximum number of iterations was reached before the traditional method converged to the optimal estimate. The method of separating variables only required 14 iterations to obtain the optimal estimate, reducing the computational time from 1128 s to 130 s. Therefore, the application of the separable nonlinear least-squares problem can improve the calculation efficiency and convergence speed of the parameter solution process. It can also provide a new method for parameter estimation in the Gaussian model for LiDAR waveform decomposition.

1. Introduction

Light detection and ranging (LiDAR) is an active remote sensing measurement system that integrates a laser altimeter, a global navigation satellite system (GNSS), and an inertial measurement unit (IMU) [1]. Currently, LiDAR is widely employed for distance measurements [2], vegetation inversion [3, 4], urban modelling [5], marine mapping [6], and other such applications. In such applications, the received waveform is mathematically described as the convolution of the emitted laser pulse and ground response function [7]. The geometry, physical characteristics, and vertical distribution of the objects irradiated by the laser pulses can be obtained by analyzing the properties of the received waveform, such as its amplitude and pulse width.

Waveform decomposition is an important step when processing full-waveform LiDAR data. The reflected pulse corresponding to each object is extracted from the received signal, using waveform modelling. Using LiDAR full-waveform data is a prerequisite for terrain and building reconstruction [8]. As LiDAR data are dependent on several factors, such as the background light and electrical noise from the avalanche photodiode, the selection of an accurate decomposition model and an efficient and fast algorithm is important to ensure accurate and precise results. For a majority of the received signals, amplitude modulation approximates a Gaussian distribution; therefore, waveform signal decomposition based on Gaussian fitting is the most popular method used for LiDAR full-waveform decomposition [9, 10]. Hofton et al. [11] screened the waveform data

according to the importance of the Gaussian component and then solved the optimal value of each waveform component parameter using the nonlinear optimization Levenberg–Marquardt (LM) algorithm. Wagner et al. [12] analyzed the theoretical basis of simulating a LiDAR waveform using a Gaussian mixture model and obtained the pulse amplitude, half-width, and distance via Gaussian decomposition. These factors were included in the calibration equation, and the data were calibrated using an external reference to estimate the target backscattering cross section. Lai et al. [13] decomposed the waveform components from the original waveform data, layer by layer, until the maximum peak value of the remaining waveform was less than a certain threshold. Thereafter, the initial parameters were optimized using the limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm, and the optimal solution of the waveform component was finally obtained. Lu et al. [14] improved the second-order reciprocal zero-point method to detect the initial position of waveform components and then iteratively optimized the initial value using the LM algorithm to achieve least-squares fitting of the waveforms.

In general, the Gaussian decomposition of waveforms consists of two steps: (1) determining the number of Gaussian components and their initial parameters and (2) optimizing and postprocessing the characteristic parameters. A majority of parameter estimation methods regard all Gaussian parameters as nonlinear variables [15] and then employ the nonlinear optimization method to solve the parameters. Liu et al. [16] combined this method with sequential quadratic programming (SQP), developing a gradient-based optimization algorithm; it determined the optimal time delays and system parameters in a novel dynamic optimization problem for nonlinear multistage systems. To address the randomness of setting the final iteration condition artificially in the process of obtaining the optimal valuation, Zhai et al. [17] proposed a novel characteristic value correction iteration method, and L curve was drawn to express the correlation of solutions and residuals and the iteration time corresponding to the maximum curvature point is chosen as the iteration termination condition to obtain the “optimal” intermediate estimates. To enhance the estimation accuracy, Ding et al. [18, 19] presented a filtering-based gradient iterative algorithm and a filtering-based least-squares iterative algorithm, which yielded an improved convergence speed. However, a majority of the classic LiDAR waveform decomposition methods focus on determining the initial parameters and the iterative algorithm of parameter optimization. These methods neglect the feature that the full-waveform model is a linear combination of multiple Gaussian components. Moreover, the traditional nonlinear optimization method relies on the initial value of all parameters involved in the calculation, which requires a high demand on the parameter optimization algorithm. If some of these initial parameters are not selected appropriately, the waveform parameters may not converge to the optimal estimation value, thereby worsening the waveform processing effect.

Golub and Pereyra proposed an explicit variable projection (VP) method for solving the separable nonlinear

least-squares problem in 1973 [20]. Compared with the corresponding algorithm, which directly solves the nonlinear minimum resultant problem, it featured three advantages: fewer iteration steps, fewer initial point guesses, and potentially reduced ill-conditioned degree when the original problem is ill conditioned. Based on this method, the improvement and application [21–23] of VP were conducted. Kaufman [24] proposed a modified VP algorithm based on trapezoidal decomposition to simplify the calculation, which reduced its computational complexity and improved computational efficiency. Subsequently, Ruano et al. [25] proposed an improved VP algorithm based on orthogonal triangle (QR) decomposition for the sparse nonlinear function matrix, which also improved operation efficiency. According to different calculation methods of the Jacobian matrix of variable projection algorithm, Gan et al. [26] summed up four separation algorithms to solve the problem of separable nonlinear least squares and proved that it is usually very beneficial to separate the variables into a linear and a nonlinear part. Furthermore, for cases where the number of observations was significantly higher than the number of linear parameters, Gan and Li [27] proposed a VP algorithm based on Gram–Schmidt matrix decomposition, which reduces the amount of calculation required. In addition, Chen et al. [28] proposed a modified Gram–Schmidt method-based variable projection algorithm for separable nonlinear models and verified its effectiveness experimentally. Wang et al. [29] proposed a variable projection algorithm based on singular value decomposition (SVD) to solve the separable nonlinear least-squares problem. SVD was adopted to simplify the VP algorithm and improve the stability of the calculation. Subsequently, after the elimination of the linear parameters by the improved VP algorithm, the separable least-squares problem was transformed into an optimization problem only containing nonlinear parameters.

In this study, we applied the variable projection algorithm based on the SVD of the separable nonlinear least-squares problem for the waveform fitting of received LiDAR signals. Linear parameters were expressed with nonlinear parameters by the function, and the problem was then transformed into a least-squares problem only containing nonlinear parameters. Finally, the improved LM algorithm [30, 31] was employed to solve the nonlinear parameters, whereas the linear parameters were solved using the variable projection algorithm for nonlinear functions.

The remainder of this paper is organized as follows. In Section 2, the method for determining the number of components and the initial values of the Gaussian decomposition parameters for LiDAR waveforms is presented, along with a brief description of the traditional method for estimating waveform parameters. Section 3 presents the proposed method for separating linear and nonlinear parameters in the waveform model based on a variable projection algorithm; the optimization algorithm and the processing of waveform parameters after separation are also discussed in this section. Subsequently, the methods with and without separation of the linear and nonlinear parameters, i.e., the proposed and existing methods,

respectively, are compared using numerical experiments, as described in Section 4. Finally, the conclusions and implications of the study are discussed in Section 5.

2. Gaussian Decomposition of LiDAR Signals

A Gaussian mixture model with N Gaussian functions was used to describe the received LiDAR waveform according to the radar range formula. After fitting and decomposing, the received pulses related to single objects are separated and extracted. Finally, the attribute information of each signal received from the ground is obtained.

Assuming that a received waveform has m sampling points and is composed of N waveform components, the waveform based on the Gaussian function model is expressed as follows:

$$f(x, \mathbf{a}, \boldsymbol{\mu}, \boldsymbol{\sigma}) = \sum_{s=1}^N a_s \exp\left(-\frac{(x - \mu_s)^2}{2\sigma_s^2}\right) + \text{noise}, \quad (1)$$

where a_s , μ_s , and σ_s are the pulse amplitude, center position, and half-width of the s^{th} waveform component, respectively. Here, noise denotes the noise in the original waveform. The aim of full-waveform Gaussian decomposition is to optimize $3N$ parameters of Gaussian components, where the number of linear parameters (a_s) is N and the number of nonlinear parameters (μ_s and σ_s) is $2N$.

2.1. Determining the Number of Gaussian Waveform Components and Parameter Initialization. To determine the components of the Gaussian waveform, the amplitude threshold is set and the first derivative of the function is calculated. When the maximum value is greater than the amplitude threshold, it is determined as the peak value. Thereafter, the second derivative of the function is calculated, and the inflection point is obtained. The time difference between two adjacent inflection points should be greater than the half-width of the transmitted pulse. Finally, the peak and inflection points are matched, where the position of the peak should lie between the two inflection points.

After determining the number of Gaussian function components, the feature parameters are initialized, where t_{2s-1} and t_{2s} are set as the sampling times of adjacent inflection points. Subsequently, the center position μ_s and the half-width σ_s of the s^{th} Gaussian component are calculated using equations (2) and (3), respectively. The pulse amplitude a_s refers to the amplitude corresponding to the center position:

$$\mu_s = \frac{(t_{2s-1} + t_{2s})}{2}, \quad (2)$$

$$\sigma_s = \frac{(t_{2s} - t_{2s-1})}{2}. \quad (3)$$

2.2. LM Optimization without Waveform Parameter Separation. After determining the number of components and initial parameters of the Gaussian function, the most

commonly used method regards both linear and nonlinear parameters as nonlinear parameters and subsequently solves them using the LM algorithm, for the nonlinear least-squares optimization [20]. The objective function is constructed by minimizing the sum of the squares of errors between the fitted and measured waveforms after establishing the Gaussian function model. According to the principle of nonlinear least squares, the solutions of linear parameters $\hat{\mathbf{a}} = (\hat{a}_1, \hat{a}_2, \dots, \hat{a}_N)^T$ and nonlinear parameters $\hat{\mathbf{b}} = (\hat{\mu}_1, \hat{\mu}_2, \dots, \hat{\mu}_N, \hat{\sigma}_1, \hat{\sigma}_2, \dots, \hat{\sigma}_N)^T$ minimize the following functions:

$$\begin{aligned} \min \|\mathbf{r}\|_2^2 &= \min \sum_{i=1}^m (y_i - f(x_i, \mathbf{a}, \boldsymbol{\mu}, \boldsymbol{\sigma}))^2 \\ &= \min \sum_{i=1}^m \left(y_i - \sum_{s=1}^N a_s \exp\left(-\frac{(x_i - \mu_s)^2}{2\sigma_s^2}\right) \right)^2 \\ &= \min \sum_{i=1}^m \left(y_i - \sum_{s=1}^N a_s \phi_s(\mu_s, \sigma_s; x_i) \right)^2, \end{aligned} \quad (4)$$

where, y_i is the original observation signal, m is the number of observations, $\phi_s(\mu_s, \sigma_s; x_i) = \exp(-\frac{(x_i - \mu_s)^2}{2\sigma_s^2})$ is the s^{th} nonlinear function component, \mathbf{r} is the residual vector, and $\|\cdot\|$ is the Euclidean norm.

The parameters to be determined are

$$\begin{aligned} \mathbf{c} &= (a_1, a_2, \dots, a_N, \mu_1, \mu_2, \dots, \mu_N, \sigma_1, \sigma_2, \dots, \sigma_N) \\ &= (c_1, c_2, \dots, c_N, c_{N+1}, c_{N+2}, \dots, c_{2N}, c_{2N+1}, c_{2N+2}, \dots, c_{3N}) \\ &= (c_1, c_2, \dots, c_{3N}). \end{aligned} \quad (5)$$

The function model is expressed as

$$f(x_i, \mathbf{c}) = \sum_{j=1}^N c_j \phi_j(c_{j+N}, c_{j+2N}, x_i), \quad i = 1, 2, \dots, m. \quad (6)$$

The objective function is established under the least-squares principle:

$$F(\mathbf{c}) = \|\mathbf{r}\|_2^2 = \sum_{i=1}^m (y_i - f(x_i, \mathbf{c}))^2. \quad (7)$$

The Jacobian matrix $\mathbf{J}(\mathbf{c})$ consisting of the first-order partial derivatives of the residual vector \mathbf{r} is

$$\mathbf{J}(\mathbf{c}) = \begin{bmatrix} \frac{\partial r(x_1, \mathbf{c})}{\partial c_1} & \frac{\partial r(x_1, \mathbf{c})}{\partial c_2} & \dots & \frac{\partial r(x_1, \mathbf{c})}{\partial c_{3N}} \\ \frac{\partial r(x_2, \mathbf{c})}{\partial c_1} & \frac{\partial r(x_2, \mathbf{c})}{\partial c_2} & \dots & \frac{\partial r(x_2, \mathbf{c})}{\partial c_{3N}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial r(x_m, \mathbf{c})}{\partial c_1} & \frac{\partial r(x_m, \mathbf{c})}{\partial c_2} & \dots & \frac{\partial r(x_m, \mathbf{c})}{\partial c_{3N}} \end{bmatrix}. \quad (8)$$

Thus, the Hessian matrix of \mathbf{r} is approximated as

$$\mathbf{H}(\mathbf{c}) = \mathbf{J}^T(\mathbf{c})\mathbf{J}(\mathbf{c}) + \lambda\mathbf{I}. \quad (9)$$

Therefore, the LM optimization algorithm is used to iterate the parameters:

$$\begin{aligned} \mathbf{c}^{(k+1)} &= \mathbf{c}^{(k)} + \alpha^{(k)}\mathbf{d}^{(k)} \\ &= \mathbf{c}^{(k)} - \alpha^{(k)}\left[\mathbf{H}(\mathbf{c}^{(k)}) + \lambda^{(k)}\mathbf{I}\right]^{-1}\mathbf{J}(\mathbf{c}^{(k)})\mathbf{r}(\mathbf{c}^{(k)}), \end{aligned} \quad (10)$$

where $\mathbf{d}^{(k)} = -[\mathbf{H}(\mathbf{c}^{(k)}) + \lambda^{(k)}\mathbf{I}]^{-1}\mathbf{J}(\mathbf{c}^{(k)})\mathbf{r}(\mathbf{c}^{(k)})$, $\mathbf{c}^{(k)}$ is the k^{th} iteration result, $\mathbf{d}^{(k)}$ is the iteration direction, and $\alpha^{(k)}$ is the iteration length. To ensure that the objective function, i.e., Equation (6), is always in a descending state, $\alpha^{(k)}$ is calculated using the Armijo criterion [32] of imprecise search. The term λ is adjusted using a strategy similar to the radius of the trust region [33]. First, a quadratic function is defined at the current iteration point to represent the quadratic approximation of the objective function:

$$\begin{aligned} v(\mathbf{c}^{(k)} + \mathbf{d}^{(k)}) &= F(\mathbf{c}^{(k)}) + (\mathbf{J}^T(\mathbf{c}^{(k)})\mathbf{r}^{(k)}(\mathbf{c}^{(k)}))^T\mathbf{d}^{(k)} \\ &\quad + \frac{1}{2}\mathbf{d}^{(k)T}(\mathbf{J}^T(\mathbf{c}^{(k)})\mathbf{J}(\mathbf{c}^{(k)}))\mathbf{d}^{(k)}. \end{aligned} \quad (11)$$

Thereafter, the ratio of the actual decrease in $v(\mathbf{d})$ to the expected decrease in the objective function F is calculated based on the iteration direction:

$$\begin{aligned} \eta^{(k)} &= \frac{\Delta F(\mathbf{d}^{(k)})}{\Delta v(\mathbf{d}^{(k)})} = \frac{F(\mathbf{c}^{(k)} + \mathbf{d}^{(k)}) - F(\mathbf{c}^{(k)})}{v(\mathbf{c}^{(k)} + \mathbf{d}^{(k)}) - v(\mathbf{c}^{(k)})} \\ &= \frac{F(\mathbf{c}^{(k)} + \mathbf{d}^{(k)}) - F(\mathbf{c}^{(k)})}{(\mathbf{J}^T(\mathbf{c}^{(k)})\mathbf{r}^{(k)}(\mathbf{c}^{(k)}))^T\mathbf{d}^{(k)} + 1/2(\mathbf{d}^{(k)})^T(\mathbf{J}^T(\mathbf{c}^{(k)})\mathbf{J}(\mathbf{c}^{(k)}))\mathbf{d}^{(k)}}. \end{aligned} \quad (12)$$

For an initial value of λ , $\mathbf{d}^{(k)}$ is calculated at each step in the LM algorithm. Subsequently, $\lambda^{(k)}$ is adjusted according to the value of $\eta^{(k)}$. Finally, a line search is performed to complete the iteration. When η_k is close to one, the fitting between the quadratic function $v(\mathbf{d})$ and the objective function is good at point \mathbf{x}_k . The parameter λ should be smaller when the LM algorithm is used to solve the nonlinear least-squares problem. When η_k is close to zero, the fitting between the quadratic function $v(\mathbf{d})$ and the objective function is poor at point \mathbf{x}_k . The parameter λ should be larger when LM is used to solve the nonlinear least-squares problem. When η_k is neither close to zero nor one, λ is suitable and does not require adjustment. The critical values are typically 0.25 and 0.75, and the rules of adjusting λ are as follows:

$$\lambda^{(k+1)} = \begin{cases} 0.1\lambda^{(k)}, & \eta^{(k)} > 0.75, \\ \lambda^{(k)}, & 0.25 \leq \eta^{(k)} \leq 0.75, \\ 10\lambda^{(k)}, & \eta^{(k)} < 0.25. \end{cases} \quad (13)$$

3. LM Optimization Method for Waveform Parameters Based on Variable Projection

3.1. Gaussian Model Based on a Variable Projection Algorithm. Considering the structural characteristics of the functional model for the received LiDAR signal, we derived a new equation as a linear combination of $\exp(-(x - \mu_s)^2/2\sigma_s^2)$ with respect to a_s . Equation (4) is written in the matrix form as

$$\min \|\mathbf{r}\|^2 = \min \|\mathbf{y} - \Phi(\mathbf{b})\mathbf{a}\|^2, \quad (14)$$

where \mathbf{y} is the vector of the original signal. The column of $\Phi(\mathbf{b})$ corresponding to the nonlinear function is $\phi_k(\mathbf{b}; \mathbf{x})$ for parameter \mathbf{b} . For the given nonlinear parameters \mathbf{b} , the linear parameter \mathbf{a} can be estimated by solving the following linear least-squares problem:

$$\mathbf{a} = (\Phi^T(\mathbf{b})\Phi(\mathbf{b}))^{-1}\Phi(\mathbf{b})\mathbf{y} = \Phi^+(\mathbf{b})\mathbf{y}, \quad (15)$$

where $\Phi^+(\mathbf{b})$ is the pseudoinverse of $\Phi(\mathbf{b})$. Substituting (20) in (19) yields

$$\begin{aligned} \min \|\mathbf{r}_2\|^2 &= \min \|\mathbf{y} - \Phi(\mathbf{b})\Phi^+(\mathbf{b})\mathbf{y}\|^2 \\ &= \min \|(\mathbf{I} - \Phi(\mathbf{b})\Phi^+(\mathbf{b}))\mathbf{y}\|^2 \\ &= \min \|(\mathbf{I} - \mathbf{P}_{\Phi(\mathbf{b})})\mathbf{y}\|^2 \\ &= \min \|\mathbf{P}_{\Phi(\mathbf{b})}^\perp\mathbf{y}\|^2, \end{aligned} \quad (16)$$

where $\mathbf{P}_{\Phi(\mathbf{b})}$ is the orthogonal projector on the linear space spanned by the columns of $\Phi(\mathbf{b})$ and $\mathbf{P}_{\Phi(\mathbf{b})}^\perp$ is the projector on the orthogonal complement of vector \mathbf{y} in the column space. To simplify the calculation, the matrix composed of nonlinear functions is decomposed using SVD:

$$\Phi(\mathbf{b}) = \mathbf{U}\mathbf{S}\mathbf{V}^T, \quad (17)$$

where \mathbf{U} is an $m \times m$ orthogonal matrix, \mathbf{S} is an $m \times p$ diagonal matrix, and \mathbf{V} is a $p \times p$ orthogonal matrix. We obtained $\Phi(\mathbf{b})$ as

$$\Phi^+(\mathbf{b}) = \mathbf{V}\mathbf{S}^{-1}\mathbf{U}^T. \quad (18)$$

Thus,

$$\mathbf{P}_{\Phi(\mathbf{b})} = \Phi(\mathbf{b})\Phi(\mathbf{b}) = \mathbf{U}\mathbf{S}\mathbf{V}^T\mathbf{V}\mathbf{S}^{-1}\mathbf{U}^T = \mathbf{U} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \mathbf{U}^T. \quad (19)$$

If the rank of the matrix $\Phi(\mathbf{b})$ is r , the first r elements on the diagonal of \mathbf{S} are nonzero, i.e., $r \leq p$. Therefore, \mathbf{U} can be divided into an $m \times r$ matrix \mathbf{U}_1 and an $m \times (m - r)$ matrix \mathbf{U}_2 and \mathbf{V} can be divided into a $p \times r$ matrix \mathbf{V}_1 and a $p \times (p - r)$ matrix \mathbf{V}_2 :

$$\Phi(\mathbf{b}) = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \\ m \times r & m \times (m-r) \end{bmatrix} \begin{bmatrix} \mathbf{S}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^T \\ r \times p \\ \mathbf{V}_2^T \\ (p-r) \times p \end{bmatrix}, \quad (20)$$

$$\mathbf{P}_{\Phi(\mathbf{b})}^\perp = \mathbf{I}_{m \times m} - \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \\ m \times r & m \times (m-r) \end{bmatrix} \begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U}_1^T \\ r \times m \\ \mathbf{U}_2^T \\ (m-r) \times m \end{bmatrix} = - \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \\ m \times r & m \times (m-r) \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{p-r} \end{bmatrix} \begin{bmatrix} \mathbf{U}_1^T \\ r \times m \\ \mathbf{U}_2^T \\ (m-r) \times m \end{bmatrix} = \mathbf{U}_2 \mathbf{U}_2^T. \quad (21)$$

Therefore, the minimization problem of the objective function (15) is simplified as

$$\min \|\mathbf{r}_3\|^2 = \min \|\mathbf{U}_2 \mathbf{U}_2^T \mathbf{y}\|^2. \quad (22)$$

3.2. Optimal Estimation Method for Waveform Gaussian Parameters. To fit the separated LiDAR signal with the Gaussian decomposition model, we combined the differential equation of the pseudoinverse matrix and the variable projection algorithm. The Jacobian matrix [20] of the separated objective function is

$$\begin{aligned} \mathbf{J}(\mathbf{b}) &= \mathbf{D}(\mathbf{I} - \Phi(\mathbf{b})\Phi^+(\mathbf{b})\mathbf{y}) \\ &= -\mathbf{D}(\Phi(\mathbf{b})\Phi^+(\mathbf{b})\mathbf{y}) \\ &= -(\mathbf{I} - \Phi(\mathbf{b})\Phi^+(\mathbf{b}))\mathbf{D}(\Phi(\mathbf{b})\Phi^+(\mathbf{b})) \\ &= -\mathbf{U}_2 \mathbf{U}_2^T \mathbf{D}(\Phi(\mathbf{b})) \mathbf{D}(\Phi(\mathbf{b})) \mathbf{V}_1 \mathbf{S}^{-1} \mathbf{U}_1^T. \end{aligned} \quad (23)$$

The objective function only containing nonlinear parameters is optimized using the LM algorithm. To ensure global optimization, once the nonlinear parameters are calculated for a given iteration, the linear parameters are calculated using equation (15) until the overall gradient is less than a threshold value. For the convenience of expression, in the following text, the traditional LM method without the separation of parameters is referred to as $\mathbf{LM}_{\text{unSep}}$. Moreover, the separable nonlinear least-squares solution method (i.e., LM method with the SVD variable projection algorithm for parameter separation) is referred to as \mathbf{LM}_{VP} . The optimization process for the waveform decomposition parameters based on the \mathbf{LM}_{VP} method is summarized in Algorithm 1.

4. Numerical Experiment

The nonlinear optimization LM method for parameter separation proposed herein for the estimation of Gaussian decomposition parameters of LiDAR signals was validated using five groups of simulated LiDAR signals and the ICESat/GLAS satellite LiDAR waveform data for the terrain of Germany in 2003. The same data were evaluated using the traditional nonlinear optimization method to compare the two models in terms of solution results, evaluation coefficient, and convergence. The experiments were performed using MATLAB 2016b on a 2.3 GHz desktop PC running Windows 10.

4.1. Example 1: Simulation Experiment. It was assumed that the five groups of simulated received LiDAR signals were composed of four Gaussian components, with a sampling interval of 0.5 ns and a sampling number of 200. The amplitude, position, and width of the original four Gaussian functions in the five groups are listed in Table 1.

Due to the influence of noise, the actual waveform distribution was not completely Gaussian. Based on the simulation data, random noise obeying a standard deviation of 0.5 was added to obtain the final five groups of discrete received signals. First, according to the parameter initialization method, the initial values were selected. Thereafter, the waveform parameters were iteratively calculated using both methods for a comparison.

4.1.1. Comparisons of Solution Results in Simulation Experiment. For the five groups of data, the results of the optimization methods with and without parameter separation are listed in Table 2.

From Table 2, it is evident that the results for both methods were identical, except in the case of group 4. For a clearer visualization of the decomposition and fitting effect, the original observation value, decomposition waveform, and fitting waveform obtained using both methods for the five groups of signals are depicted in Figures 1 and 2. As the fitting waveform was identical to the decomposition waveform obtained using the $\mathbf{LM}_{\text{unSep}}$ method and the \mathbf{LM}_{VP} method, except in the case of group 4, groups 1, 2, 3, and 5 were represented as a graph, as depicted in Figure 1. Figures 2(a) and 2(b) present the results for group 4 obtained using the $\mathbf{LM}_{\text{unSep}}$ method and the \mathbf{LM}_{VP} method, respectively.

As shown in Figure 1, the decomposition waveforms calculated via the two methods are identical. In Figure 2, the decomposition waveforms and the fitting waveforms obtained for group 4 are inconsistent. After comparing Figures 2(a) and 2(b) and the parameter results in Table 2, we observe that the optimization method and the decomposition waveform based on variable projection separation fit the received signal more closely than those of the traditional optimization method. Moreover, when the iterations reach the maximum limit, the traditional parameter optimization method did not converge on the optimal value.

4.1.2. Analysis of Simulation Example Results. To analyze the effect of the parameter optimization method based on variable projection separation, we compared it with the

Step 1: for initial values for $\mathbf{a}^{(0)} \in \mathfrak{R}^N$, $\mathbf{b}^{(0)} \in \mathfrak{R}^{2N}$, set $\lambda^{(0)} = 0.5$, threshold value $\varepsilon = 10^{-6}$, and $k = 1$; the maximum number of iterations is 100.

Step 2: compute $\Phi(\mathbf{b}^{(k)})$ and $\mathbf{J}^{(k)}(\mathbf{b})$ using equation (14) and (23), respectively.

Step 3: compute the iteration length $\alpha^{(k)}$ using the Armijo criterion and the iteration direction $\mathbf{d}^{(k)}$.

Step 4: adjust $\lambda^{(k)}$ according to $\eta^{(k)}$ and then compute $\mathbf{b}^{(k)}$ using the LM algorithm iteration, i.e., equation (10).

Step 5: compute $\mathbf{a}^{(k)}$ according to the variable projection algorithm.

Step 6: if $\|\mathbf{g}^{(k)}\|_2 = \|(\mathbf{J}^{(k)}(\mathbf{b}))^T \mathbf{r}_3\|_2 \leq \varepsilon$, stop. Else, $k = k + 1$ and go to step (3).

ALGORITHM 1: Parameter optimization process of the LM_{VP} method.

TABLE 1: Five groups of original parameters of Gaussian function components.

Group	a	μ	σ
1	55, 45, 50, 20	25, 40, 55, 85	4, 4.5, 4, 3.5
2	39, 45, 39, 80	19, 30, 60, 80	4, 4.5, 4, 3.5
3	39, 50, 80, 60	19, 30, 60, 75	4, 4.5, 4, 3.5
4	45, 39, 80, 35	19, 30, 42, 55	4, 4.5, 4, 3.5
5	70, 40, 45, 39	15, 45, 57, 78	4, 4, 3.5, 4.5

TABLE 2: Simulation results for the five groups of data calculated using the two methods without separation (LM_{unSep}) and with separation (LM_{VP}) of parameters.

Par	Meth									
	1		2		3		4		5	
	LM _{unSep}	LM _{VP}								
a_1	55.0617	55.0617	38.7165	38.7165	39.1848	39.1848	45.4493	44.5576	69.9570	69.9570
a_2	44.6495	44.6495	44.9162	44.9162	50.0122	50.0122	30.1706	39.0140	39.8209	39.8209
a_3	50.0565	50.0565	38.8597	38.8597	79.9920	79.9920	69.6504	79.5206	44.7605	44.7605
a_4	20.0493	20.0493	79.9784	79.9784	59.7621	59.7621	23.4229	35.2598	38.9916	38.9916
μ_1	24.9878	24.9878	19.0105	19.0105	19.0048	19.0048	22.4230	18.8942	15.0032	15.0032
μ_2	3.9835	3.9835	4.0028	4.0028	3.9963	3.9963	6.8918	3.9122	4.0109	4.0109
μ_3	39.9723	39.9723	29.9902	29.9902	30.0025	30.0025	31.7396	29.9838	44.9793	44.9793
μ_4	4.5465	4.5465	4.5212	4.5212	4.4845	4.4845	-0.1427	4.7580	3.9749	3.9749
σ_1	55.0010	55.0010	59.9848	59.9848	59.9971	59.9971	41.6845	42.0733	56.9878	56.9878
σ_2	3.9996	3.9996	3.9960	3.9960	3.9939	3.9939	5.5996	3.9718	3.5163	3.5163
σ_3	85.0168	85.0168	79.9924	79.9924	74.9912	74.9912	55.9326	55.0211	77.9833	77.9833
σ_4	3.4960	3.4960	3.4979	3.4979	3.5066	3.5066	3.3875	3.4721	4.5122	4.5122

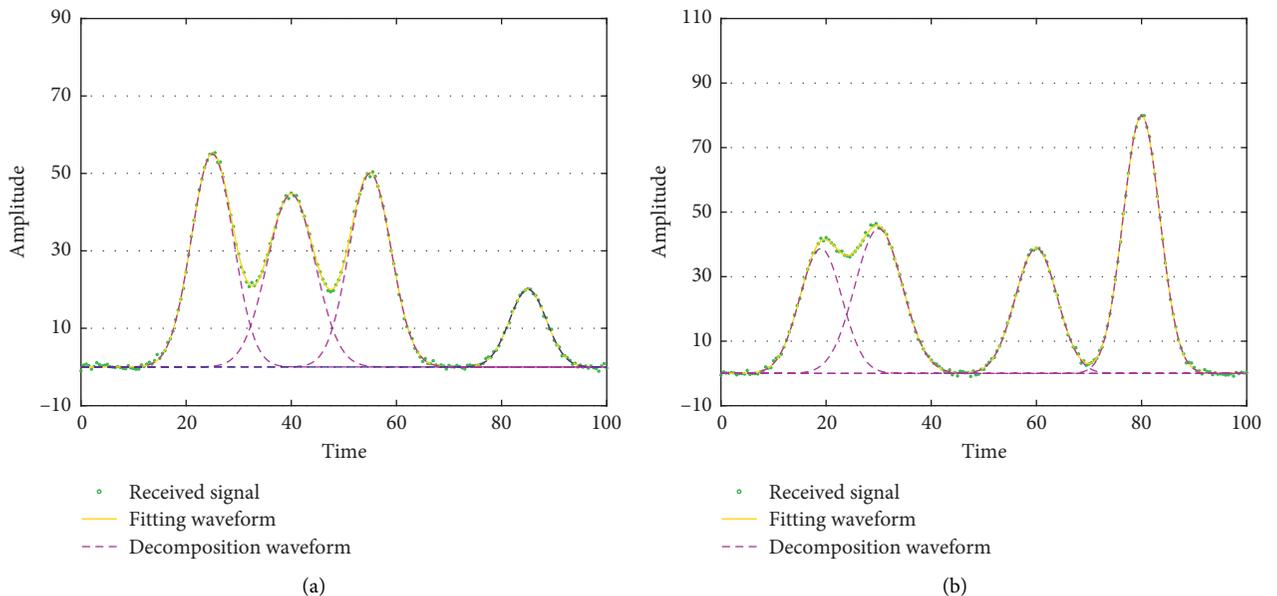


FIGURE 1: Continued.

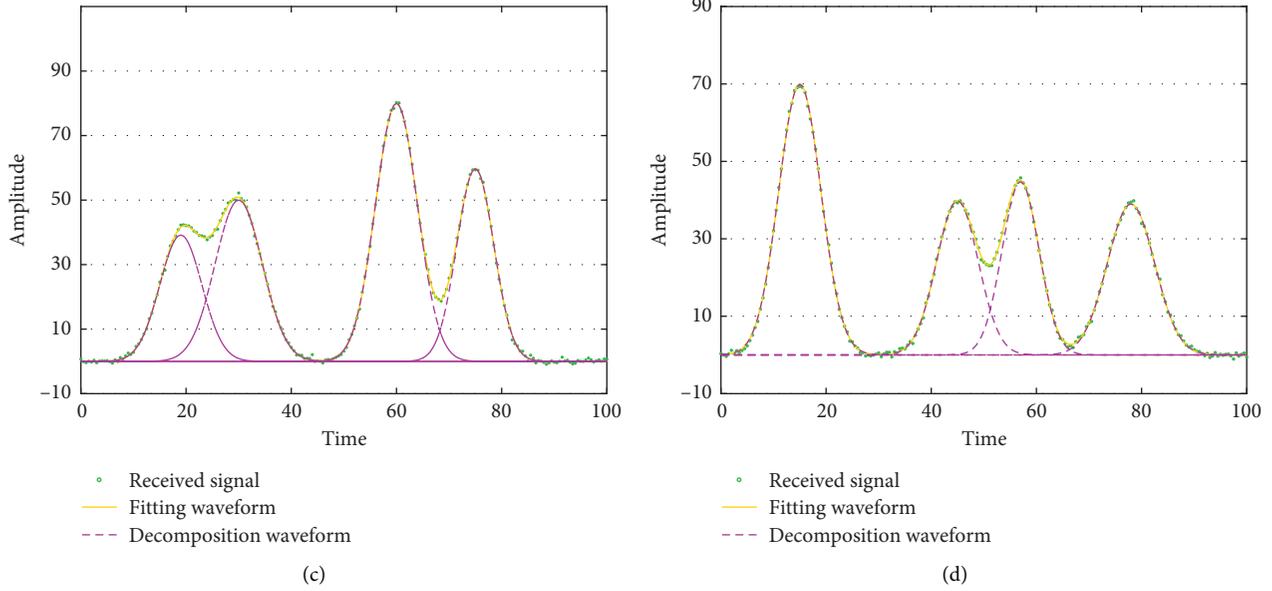


FIGURE 1: Decomposed waveforms obtained for groups (a) 1, (b) 2, (c) 3, and (d) 5.

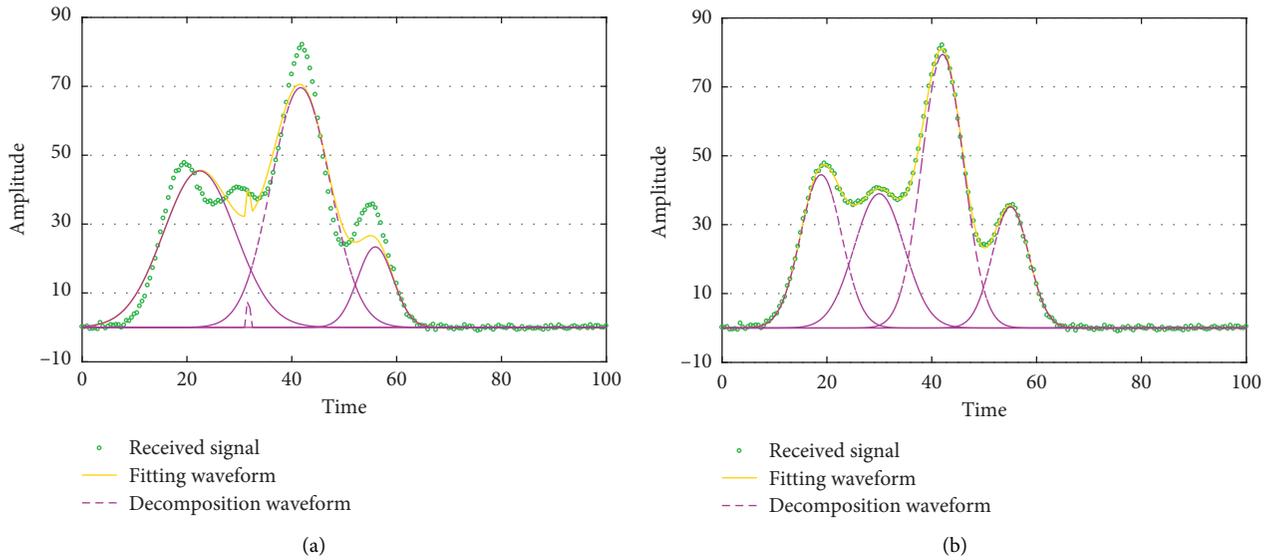


FIGURE 2: Decomposed waveforms obtained for the two optimization methods (a) without and (b) with parameter separation for group 4.

traditional parameter optimization method in terms of iterations (Iterations), computing time (T , the unit is in seconds), root mean square error (RMSE), goodness of fit (R^2), correlation coefficient (C), and maximum deviation (M_diff):

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y'_i - y_i)^2}, \quad (24)$$

where y_i is the original received signal and y'_i is the fitted received signal.

$$R^2 = 1 - \frac{\sum_{i=1}^m (y'_i - y_i)^2}{\sum_{i=1}^m (y_i - \bar{y}_i)^2}, \quad (25)$$

where \bar{y}_i is the average of y_i .

$$C = \frac{\sum_{i=1}^m (y'_i - \bar{y}'_i)(y_i - \bar{y}_i)}{\sqrt{\sum_{i=1}^m (y'_i - \bar{y}'_i)^2 \sum_{i=1}^m (y_i - \bar{y}_i)^2}}, \quad (26)$$

where \bar{y}'_i is the mean of y'_i . RMSE values closer to 0 indicate a better fit, whereas R^2 and C values closer to 1 indicate a better correlation between the fitting result and the original signal.

$$M_diff = \max(|y_i - y'_i|), \quad (27)$$

TABLE 3: Comparison of the evaluation coefficients obtained from analysis of the results of the two methods used to simulate the data from the five different groups.

Group	Method	Iterations	T (s)	RMSE	R^2	C	M_diff
1	LM _{unSep}	30	105.4666	0.4703	0.9993	0.9996	1.1033
	LM _{VP}	11	97.43692	0.4703	0.9993	0.9996	1.1033
2	LM _{unSep}	55	242.3767	0.4801	0.9995	0.9997	1.2654
	LM _{VP}	11	96.9239	0.4801	0.9995	0.9997	1.2654
3	LM _{unSep}	72	370.6508	0.5045	0.9995	0.9998	1.6350
	LM _{VP}	10	99.8441	0.5045	0.9995	0.9998	1.6350
4	LM _{unSep}	100	378.713	4.1704	0.9671	0.9843	8.3718
	LM _{VP}	13	128.2122	0.4562	0.9996	0.9998	1.3710
5	LM _{unSep}	26	102.7408	0.4860	0.9993	0.9997	1.3203
	LM _{VP}	10	88.31994	0.4860	0.9993	0.9997	1.3203

where M_diff is the maximum absolute value of the local difference between the received signal and the fitted signal.

Subsequently, the evaluation coefficient results were compared, as shown in Table 3.

Table 3 indicates that the calculation for group 4 data using the LM_{unSep} method required the highest number of iterations (100), and its results differed significantly from the optimal estimate (as shown in Figure 2). However, the proposed improved LM_{VP} method yielded the optimal estimate after just 13 iterations. Similarly, for the other groups of data, the improved method required fewer iterations than the LM_{unSep} method, while yielding a significantly improved computational efficiency. The RMSE, R^2 , C , and M_diff values were identical for both methods and all data, except in the case of group 4, wherein both methods obtained the optimal estimation.

4.1.3. Convergence Analysis in Simulation Experiment. To further compare the results of the LM_{unSep} and LM_{VP} methods, the parameter iteration process for the five groups of data is depicted in Figure 3. Figure 3(a) presents a comparison between the iteration processes of the linear parameter a_s for the five groups using the two different methods. Figures 3(b) and 3(c) present the same comparison for the nonlinear parameters μ_s and σ_s , respectively. Both the methods converged to the optimal value; however, the LM_{VP} method required fewer iterations, while achieving faster convergence.

4.2. Example 2: Decomposition of Real Full-Waveform LiDAR Data. In the real experiment, we used the ICESat/GLAS satellite LiDAR waveform data for the terrain of Germany in 2003 (Sources: <https://nsidc.org/data/>), with a sampling interval of 1 ns and a sampling number of 544.

4.2.1. Comparisons of Solution Results in a Real Experiment. For the LM_{unSep} method, we conducted two experiments that were restricted by different maximum number of iterations, i.e., 100 and 200 iterations. The results of the parameters were calculated using the LM_{unSep} and LM_{VP} methods (Table 4).

In the case of the LM_{unSep} method, both experiments reached the maximum number of iterations; the results for 200 iterations was always closer to the LM_{VP} method than those for 100 iterations. When the LM_{unSep} method was allowed to run 200 iterations, it approached the results from the LM_{VP} method; however, it never met them. Therefore, to obtain an optimal estimation value from the LM_{unSep} method, more than 200 iterations and additional time is required.

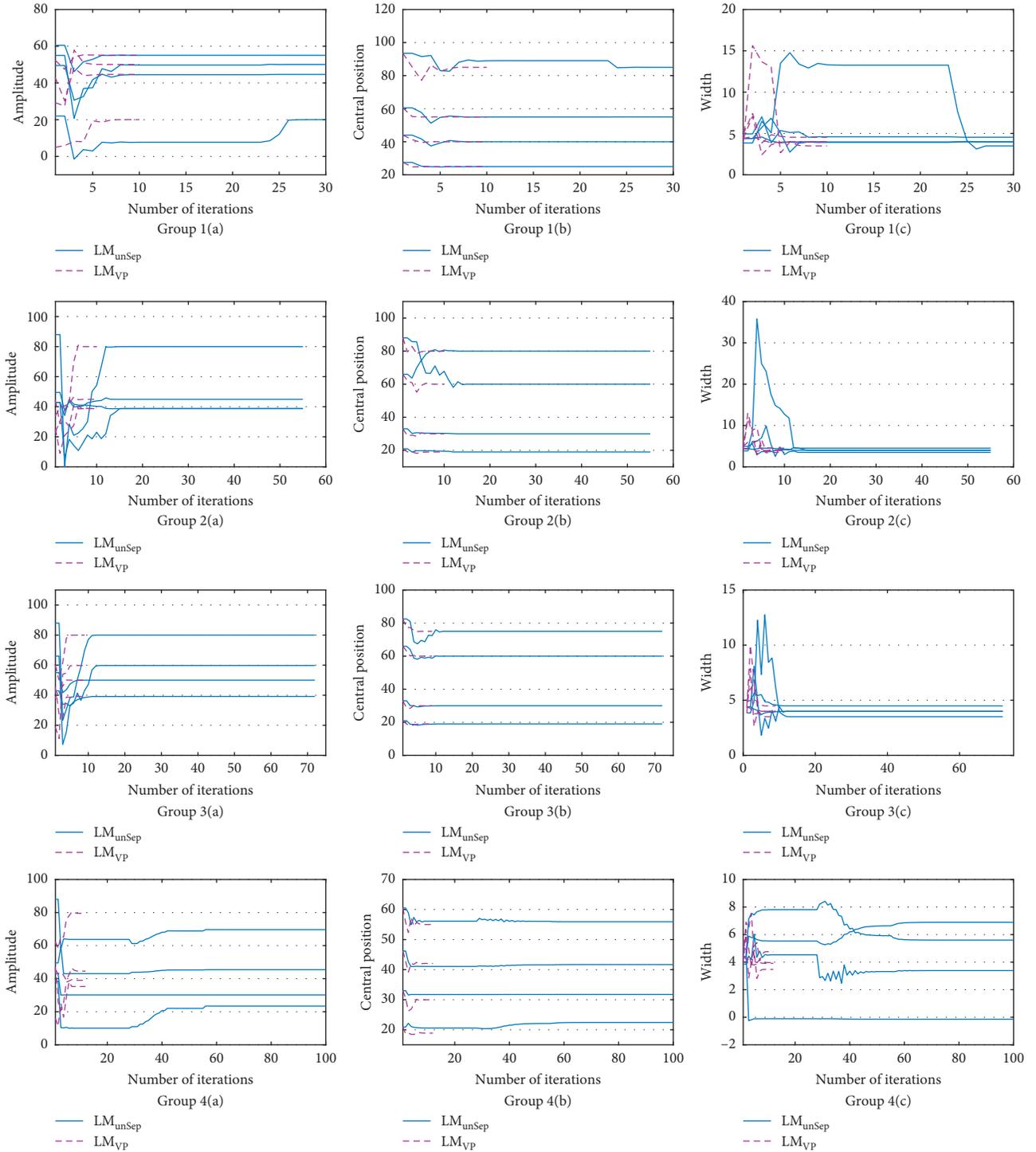
Figure 4 presents a clearer comparison between the decomposition and fitting effect for the two models, original observation value, decomposition waveform, and fitted waveform of the received LiDAR signals. Figure 4(a) presents the LM_{unSep} graph for the 200-iteration run, and Figure 4(b) depicts the LM_{VP} graph.

As in the comparison of the parameters, the fitting and decomposition waveforms for the LM_{unSep} and LM_{VP} methods yield similar results; however, the LM_{VP} method always yields the best fit.

4.2.2. Analysis of the Real Experiment Example Results. We analyzed the parameters in the LM_{unSep} and LM_{VP} method calculations to determine the evaluation coefficients (Table 5).

The two experiments using the LM_{unSep} method failed to achieve the optimal result with the maximum number of iterations; however, the LM_{VP} method achieved the optimal result within 14 iterations, i.e., a reduction of 14 times considering the 200-iteration case. The RMSE from the LM_{VP} method was the lowest, whereas the RMSE from the LM_{unSep} method using 200 instead of 100 iterations was closer to that of the LM_{VP} method. Considering the goodness-of-fit and correlation coefficient, the LM_{VP} method yielded values closer to 1 than the traditional method, indicating better fitting. Similarly, M_diff was lower for the improved method. In conclusion, in the real numerical experiment, the LM_{VP} method achieved higher accuracy and better fitting than the LM_{unSep} method.

4.2.3. Convergence Analysis in Real Experiment. The parameter iteration process depicted in Figure 5 provides a more clear comparison between the results of both models in the real experiments.



(a)

FIGURE 3: Continued.

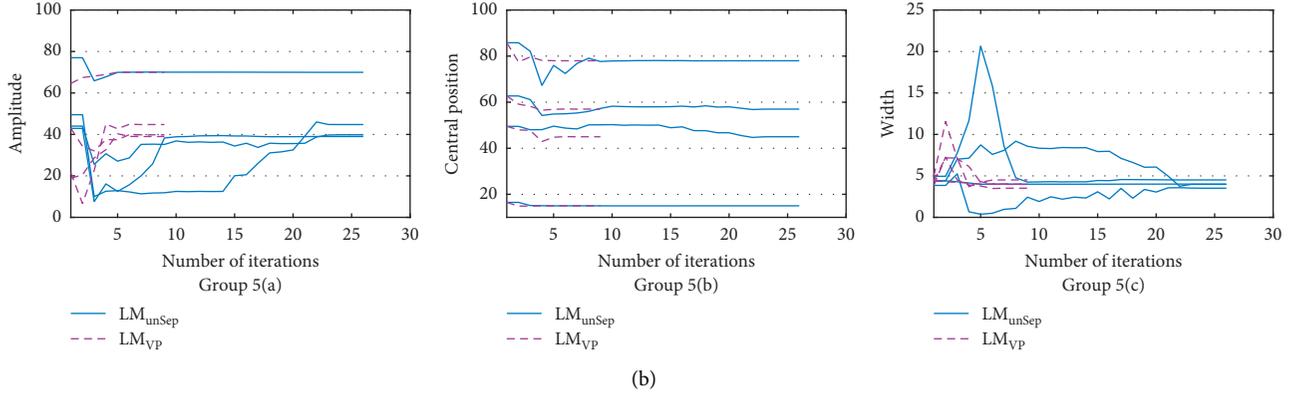


FIGURE 3: Iterative process for calculating a_s , μ_s , and σ_s from the five groups of data using the two different methods.

TABLE 4: Parameters calculated using the two models in a real experiment.

Method	Iterations	a_1	a_2	a_3	μ_1	μ_2	μ_3	σ_1	σ_2	σ_3
LM _{unSep} -100	100	0.3609	0.7340	0.4041	69.7215	90.9647	120.0394	2.9949	6.6770	9.0623
LM _{unSep} -200	200	0.3608	0.7354	0.4017	69.7300	90.9229	119.9231	3.0054	6.6211	9.2137
LM _{VP}	14	0.3606	0.7392	0.3921	69.7586	90.7592	119.4471	3.0423	6.4187	9.8273

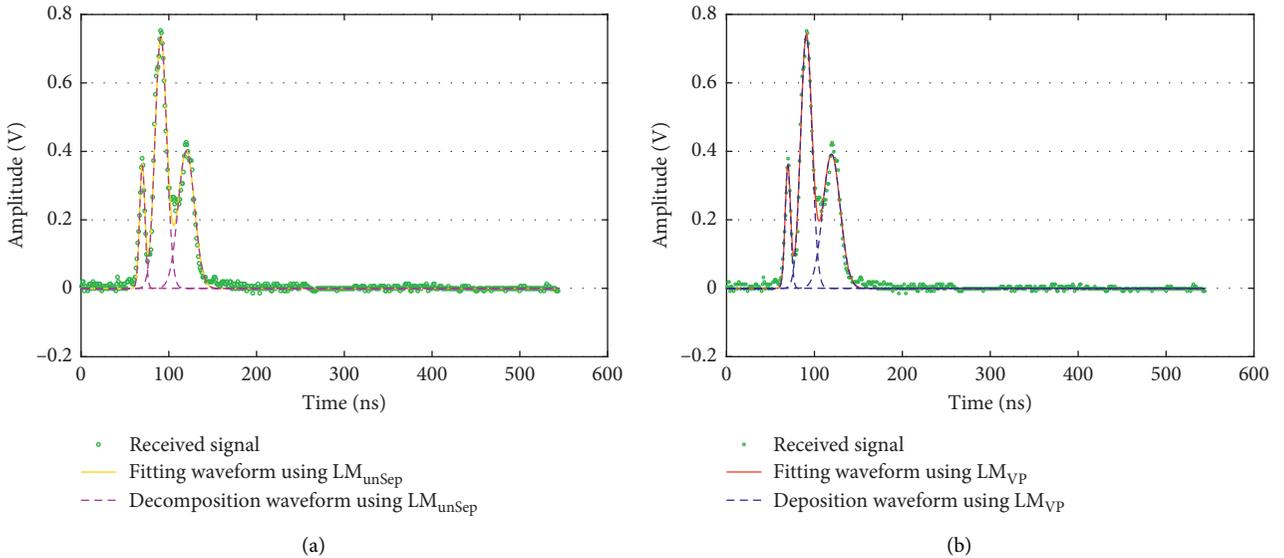


FIGURE 4: Decomposed waveforms obtained by the (a) LM_{unSep} method and the (b) LM_{VP} method for real data.

TABLE 5: Comparison of effect evaluation coefficients obtained via the two methods in the real experiment.

Method	Iterations	T (s)	RMSE	R^2	C	M_diff
LM _{unSep} -100	100	535.3555	0.01494	0.9865	0.9933	0.07899
LM _{unSep} -200	200	1128.4124	0.01479	0.9868	0.9934	0.7793
LM _{VP}	14	130.4165	0.01455	0.9872	0.9936	0.07100

Although the LM_{unSep} and LM_{VP} methods converge to a similar value, the LM_{VP} method requires fewer iterations and achieves faster convergence. For the LM_{unSep} method, there are only minor changes in the optimal value during the later iteration steps; however, for the LM_{VP} method, the

optimal value converges in 14 iterations, achieving a higher convergence speed.

Therefore, in summary, the simulation and real experiments prove that the LM_{VP} method yields better results than the LM_{unSep} method and more efficient calculations.

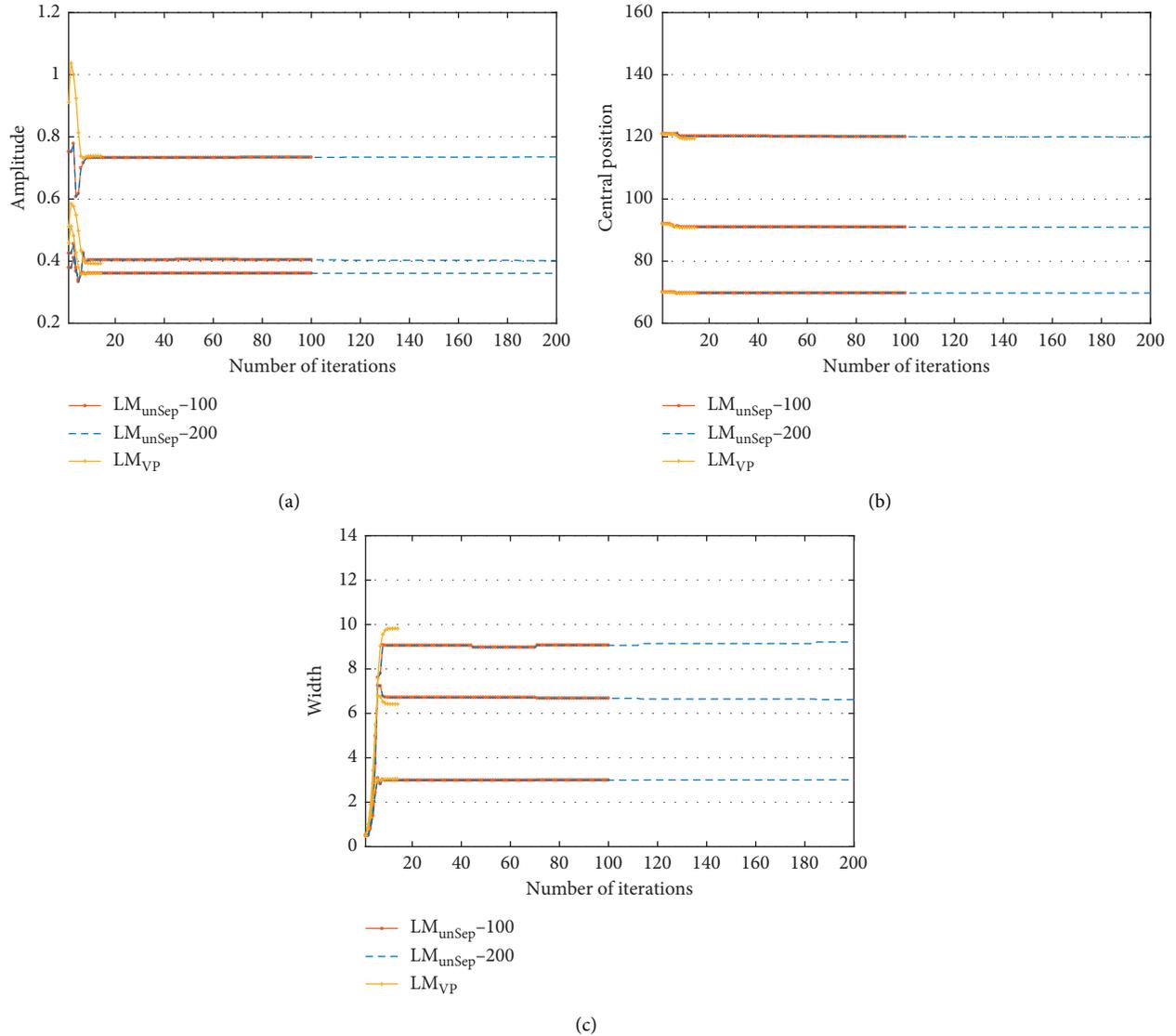


FIGURE 5: Parameter iterative process for calculating (a) a , (b) μ , and (c) σ from the satellite data in the real experiment involving the two methods.

5. Conclusions

This study focused on the special case of LiDAR signals that can be decomposed into a linear combination of multiple Gaussian function components. The linear and nonlinear parameters are separated using the SVD variable projection algorithm, and the original problem is transformed into a least-squares problem comprising only nonlinear parameters. Compared with the traditional method of parameter estimation without separation, the proposed method increased the probability of global convergence by reducing the dimension of the parameters to be estimated, especially in terms of reducing the number of iterations and computational time. In addition, the new method yielded better waveform fitting, which increased the accuracy of extracting the ground feature information from the received signal. The findings of this study are expected to expand the application of separable nonlinear least-squares methods, thereby

providing a reliable solution for parameter estimation when processing LiDAR signal data.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Nature Science Foundation of China (grant no. 41774002), the Taishan Scholars Program of Shandong Province (grant no. TSXZ201509), and

the Graduates Innovation Fund of Shandong University of Science and Technology (grant no. SDKDYC190207).

References

- [1] P. C. Li, *The Technology of Terrain and Building Reconstruction Using Airborne Full-Waveform LiDAR Data*, Information Engineering University, Zhengzhou, China, 2015.
- [2] C. E. Parrish, I. Jeong, R. D. Nowak, and R. B. Smith, "Empirical comparison of full-waveform lidar algorithms," *Photogrammetric Engineering and Remote Sensing*, vol. 77, no. 8, pp. 825–838, 2011.
- [3] M. Simard, N. Pinto, J. B. Fisher, and A. Baccini, "Mapping forest canopy height globally with spaceborne lidar," *Journal of Geophysical Research: Biogeosciences*, vol. 116, no. G4, 2011.
- [4] S. Hancock, J. Armston, Z. Li et al., "Waveform lidar over vegetation: an evaluation of inversion methods for estimating return energy," *Remote Sensing of Environment*, vol. 164, pp. 208–224, 2015.
- [5] C. Mallet, F. Bretar, M. Roux, U. Soergel, and C. Heipke, "Relevance assessment of full-waveform lidar data for urban area classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 6, pp. S71–S84, 2011.
- [6] Z. Pan, C. Glennie, P. Hartzell, J. Fernandez-Diaz, C. Legleiter, and B. Overstreet, "Performance assessment of high resolution airborne full waveform LiDAR for shallow river bathymetry," *Remote Sensing*, vol. 7, no. 5, pp. 5133–5159, 2015.
- [7] W. Wagner, "Radiometric calibration of small-footprint full-waveform airborne laser scanner measurements: basic physical concepts," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 65, no. 6, pp. 505–513, 2010.
- [8] P. J. Hartzell, C. L. Glennie, and D. C. Finnegan, "Empirical waveform decomposition and radiometric calibration of a terrestrial full-waveform laser scanner," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 1, pp. 162–172, 2015.
- [9] T. Halldórsson and J. Langerholc, "Geometrical form factors for the lidar function," *Applied Optics*, vol. 17, no. 2, pp. 240–244, 1978.
- [10] H. C. Ma, W. W. Zhou, L. Zhang, and S. Y. Wang, "Decomposition of small-footprint full waveform LiDAR data based on generalized Gaussian model and grouping LM optimization," *Measurement Science and Technology*, vol. 28, no. 4, Article ID 045203, 2017.
- [11] M. A. Hofton, J. B. Minster, and J. B. Blair, "Decomposition of laser altimeter waveforms," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 4, pp. 1989–1996, 2000.
- [12] W. Wagner, A. Ullrich, V. Ducic, T. Melzer, and N. Studnicka, "Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 60, no. 2, pp. 100–112, 2006.
- [13] X.-D. Lai, N.-N. Qin, X.-S. Han, J.-H. Wang, and W.-G. Hou, "Iterative decomposition method for small foot-print LiDAR waveform," *Journal of Infrared and Millimeter Waves*, vol. 32, no. 4, pp. 319–324, 2013.
- [14] H. Lu, Y. Pang, G. C. Xu, and Z. Y. Li, "Quantitative analysis of differences between full waveform data and system point cloud data from airborne LiDAR," *Geomatics and Information Science of Wuhan University*, vol. 40, no. 5, pp. 588–593, 2015.
- [15] Z. Q. Fu, G. L. Liu, and L. L. Guo, "Sequential quadratic programming method for nonlinear least squares estimation and its application," *Mathematical Problems in Engineering*, vol. 2019, Article ID 3087949, 8 pages, 2019.
- [16] C. Liu, R. Loxton, and K. L. Teo, "Optimal parameter selection for nonlinear multistage systems with time-delays," *Computational Optimization and Applications*, vol. 59, no. 1–2, pp. 285–306, 2014.
- [17] M. Zhai, G. Liu, Q. Tao, and M. Xin, "A novel characteristic value correction iteration method," *Communications in Statistics-Simulation and Computation*, vol. 48, no. 2, pp. 591–600, 2019.
- [18] F. Ding, L. Xu, F. E. Alsaadi, and T. Hayat, "Iterative parameter identification for pseudo-linear systems with ARMA noise using the filtering technique," *IET Control Theory and Applications*, vol. 12, no. 7, pp. 892–899, 2018.
- [19] F. Ding, F. Wang, L. Xu, and M. Wu, "Decomposition based least squares iterative identification algorithm for multivariate pseudo-linear ARMA systems using the data filtering," *Journal of the Franklin Institute*, vol. 354, no. 3, pp. 1321–1339, 2017.
- [20] G. H. Golub and V. Pereyra, "The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate," *SIAM Journal on Numerical Analysis*, vol. 10, no. 2, pp. 413–432, 1973.
- [21] Z. Q. Fu and L. L. Guo, "Tikhonov regularized variable projection algorithms for separable nonlinear least squares problems," *Complexity*, vol. 2019, Article ID 4861708, 9 pages, 2019.
- [22] D. P. O'Leary and B. W. Rust, "Variable projection for nonlinear least squares problems," *Computational Optimization and Applications*, vol. 54, no. 3, pp. 579–593, 2013.
- [23] G. Y. Chen, M. Gan, C. L. P. Chen, and H. X. Li, "A regularized variable projection algorithm for separable nonlinear least squares problems," *IEEE Transactions on Automatic Control*, vol. 64, no. 2, pp. 526–537, 2019.
- [24] L. Kaufman, "A variable projection method for solving separable nonlinear least squares problems," *BIT*, vol. 15, no. 1, pp. 49–57, 1975.
- [25] A. E. B. Ruano, D. I. Jones, and P. J. Fleming, "A new formulation of the learning problem of a neural network controller," in *Proceedings of the 30th IEEE Conference on Decision and Control*, pp. 865–866, Brighton, UK, December 1991.
- [26] M. Gan, C. L. P. Chen, G.-Y. Chen, and L. Chen, "On some separated algorithms for separable nonlinear least squares problems," *IEEE Transactions on Cybernetics*, vol. 48, no. 10, pp. 2866–2874, 2018.
- [27] M. Gan and H.-X. Li, "An efficient variable projection formulation for separable nonlinear least squares problems," *IEEE Transactions on Cybernetics*, vol. 44, no. 5, pp. 707–711, 2014.
- [28] G.-Y. Chen, M. Gan, F. Ding, and C. L. P. Chen, "Modified gram-Schmidt method-based variable projection algorithm for separable nonlinear models," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 8, pp. 2410–2418, 2019.
- [29] K. Wang, G. L. Liu, Q. X. Tao, and M. Zhai, "Efficient parameters estimation method for the separable nonlinear least squares problem," *Complexity*, vol. 2020, Article ID 9619427, 16 pages, 2020.
- [30] N. M. Nawi, A. Khan, and M. Z. Rehman, "A new Cuckoo search based Levenberg-Marquardt (CSLM) algorithm," in *Proceedings of the 13th International Conference on Computational Science and its Applications-ICCSA*, Ho Chi Minh City, Vietnam, June 2013.
- [31] C. Liu, Z. Gong, and K. L. Teo, "Robust parameter estimation for nonlinear multistage time-delay systems with noisy

- measurement data,” *Applied Mathematical Modelling*, vol. 53, pp. 353–368, 2018.
- [32] Y.-H. Li and V. Cevher, “Convergence of the exponentiated gradient method with Armijo line search,” *Journal of Optimization Theory and Applications*, vol. 181, no. 2, pp. 588–607, 2019.
- [33] Y.-X. Yuan, “Recent advances in trust region algorithms,” *Mathematical Programming*, vol. 151, no. 1, pp. 249–281, 2015.