IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received June 6, 2021, accepted June 15, 2021, date of publication June 18, 2021, date of current version June 30, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3090468

A Characterization Method of Thin Film Parameters Based on Adaptive Differential Evolution Algorithm

LIHUA LEI[1,2], (Member, IEEE), YUQING GUAN[1,2], YANHUA ZENG[1], LIN ZHAO[3], ZHANGNING XIE[1,2], ZHIGUO HAN[3], CHENGMING CAO[1], AND YUNXIA FU[1,2]

[1]Shanghai Institute of Measurement and Testing Technology, Shanghai 201203, China
[2]Shanghai Key Laboratory of Online Test and Control Technology, Shanghai 201203, China
[3]The 13th Institute of CETC, Shijiazhuang 050051, China

Corresponding author: Yunxia Fu (fuyx@simt.com)

This work was supported in part by the National Key Research and Development Program under Grant 2019YFB2005503, in part by the National Science and Technology Project of the State Administration for Market Regulation, China, under Grant 2019MK016, and in part by the Natural Science Foundation of Shanghai under Grant 21ZR1483100.

ABSTRACT According to the transmission mode of polarized light in Mueller ellipsometry, a characterization method for the thickness and optical constants of isotropic nano films based on Self-Adaptive Differential Evolution algorithm (SADE) is proposed. By establishing the least square model of the output light intensity with respect to the Mueller matrix of the standard sample to be measured, the elements of the Mueller matrix are solved by using the Sade algorithm, and the Mueller spectral curve obtained by fitting is compared with that measured by DRC-MME, and the film thickness is calculated by using the transfer matrix. The SiO_2 / Si standard samples with calibration values of 100.4 nm and 121.56 nm are simulated and calculated. The experiment shows that: when numbers of iterations accumulated to 65 and 80 respectively, the residual square sum of the light intensity of the objective function converges to the minimum values of 1.24 and 1.02. The calculated film thickness values are 101.25nm and 120.53nm respectively, and the relative errors are both less than 1%. It proves the characteristics of simple calculation, fast convergence and accurate global optimal solution.

INDEX TERMS Nanotechnology, ellipsometry, algorithms, calibration.

I. INTRODUCTION

With the development of nanometrology technology, microfilms and nanofilms have been widely used in various applications due to their excellent chemical properties. In addition to their use in microelectronics, nanofilms are used as the stealth coating of an aircraft surface and the coating on the exhaust nozzle of an aeroengine [1]–[3]. Microfilms and nanofilms typically contain monocrystalline silicon, polycrystalline silicon, silicon oxide, silicon nitride, and some metallic materials [4]. Compared with traditional coatings, microfilm and nanofilm coatings have higher hardness, oxidation resistance, and corrosion resistance. Microfilms and nanofilms improve the performance and service life of the coated material considerably [5]. The properties of nanostructured materials are affected by their optical constants, geometric dimensions, morphology, and structure. Even small changes in these parameters cause considerable variations in nanostructured material properties. The optical constant and film thickness are the most important factors influencing the properties of nanostructured materials; therefore, accurate measurement of the optical constant and thickness of a thin film is crucial [6].

Scanning electron microscopy and atomic force microscopy are the main methods of measuring the nanostructure geometry. These traditional measurement methods can meet the measurement requirements for nanostructures; however, they have the disadvantages of low measurement speed and complex operation and cannot be used to achieve real-time measurement in process manufacturing [7], [8]. The Mueller ellipsometer is a model-based optical precision measurement instrument. Geometric parameters are obtained through the Mueller matrix inversion of samples. This method has the advantages of high speed, low cost, and

The associate editor coordinating the review of this manuscript and approving it for publication was Abdullah Iliyasu.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

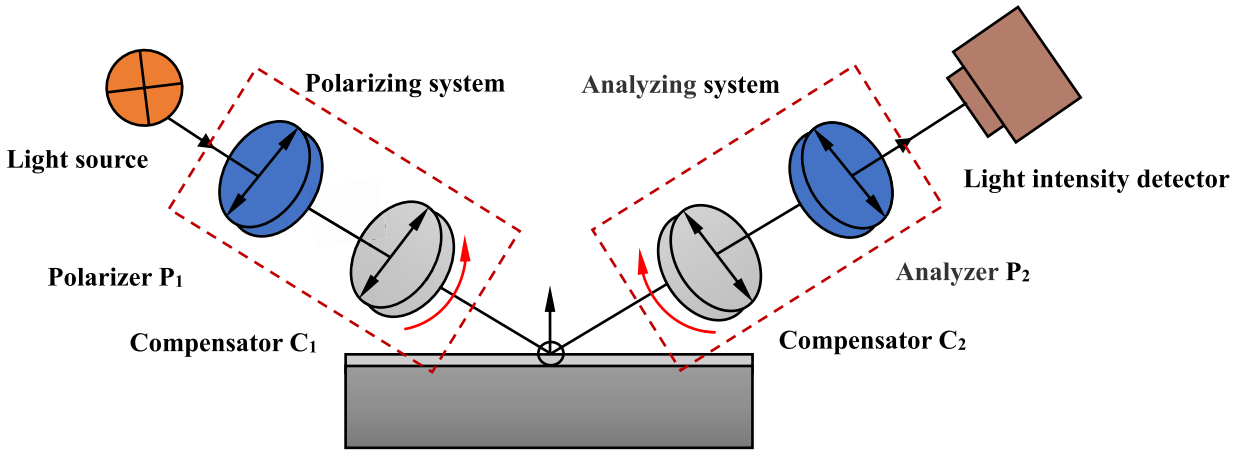

FIGURE 1. Light path diagram of light beam in ellipsoidal system.

nondestructive and simple operation and can be used to achieve the on-line detection of thin films. In general, when many samples must be measured, the solution equations related to the film parameters are overdetermined equations in modeling and calculation. Therefore, the best value in the range of multiple solution values must be found. Consequently, a strict solution cannot be obtained for the aforementioned equations [9], [10].

Levenberg-Marquardt (LM) is the main algorithm to solve the overdetermined equation of film parameters [11]. Although it combines the characteristics of Newton's method [12] and Gradient Descent (GD) [13], when the film with complex surface structure is encountered, its calculation speed is slow and accuracy is inaccurate, which does not meet the requirements of online detection in production and manufacturing. In order to discuss the requirement of high precision on-line thin film measurement, this paper proposes a method based on the self-adaptive differential evolution (SADE) algorithm for characterizing the thickness and optical parameters of nanofilms and microfilms. A least squares model of light intensity is established according to the transmission mode of a light beam in the Mueller ellipsometry system. The elements of the Mueller matrix of the sample are fitted using the SADE algorithm. Moreover, the film parameters that satisfy all the equations associated with the transmission matrix of an isotropic sample can be quickly determined with the aforementioned algorithm [14].

In this paper, the chapter I discusses the development of nanotechnology and nanostructure measurement methods. In the chapter II, the optical principle and model solving method of the double rotary spindle ellipsometry system are discussed. In the chapter III, the least square model of the outgoing light intensity is established, and the improved adaptive differential evolution algorithm is used to solve the model. In the chapter IV, experiments are designed to verify the proposed method. And we summarize the whole text in the chapter V.

II. MEASUREMENT PRINCIPLE OF AN ELLIPSOMETER

An ellipsometer measures the change in the polarization state of a light beam after it undergoes scattering through a sample. The film thickness and optical properties of the measured sample are inversely determined through modeling and calculation. The measurement process comprises two steps. The first step involves constructing a theoretical light intensity matrix comprising the geometric parameters (such as the film thickness), incident angle of the light source, and azimuth angle of the polarizer through forward optical modeling. The second step involves measuring the light intensity matrix of the sample by using the Mueller ellipsometry system and matching the theoretical optical matrix calculated by the model with the measured light intensity matrix to extract the geometric parameters when the accuracy requirement is satisfied. The light path of a beam in an ellipsometer is displayed in FIGURE 1 [1].

The Mueller ellipsometer mainly consists of a light source, polarizer, wave plate, and spectral detector [15]. The light beam emitted from the light source passes through the collimating system to form parallel light, which is vertically incident into a polarization state generator (PSG) composed of a polarizer P and compensator C_1, and the parallel light is modulated into polarized light. The polarized light then converges on the sample surface and enters a polarization state analyzer (PSA) after being reflected by the sample. The PSA is composed of a compensator C_2 and polarizer A. After demodulation by the PSA system, the emitted parallel light enters the spectrometer, which splits the light and measures the light intensity through a detector [16].

The polarization state of light is described by the Stokes vector, and the change in the polarization state of outgoing and incident light can be expressed as follows [17]:

$$S_{out} = M_A R(A_S) \cdot [R(-C_{S2}) M_{C2}(\delta_2) R(C_{S2})] \cdot M_L \cdot [R(-C_{S1}) M_{C1}(\delta_1) R(C_{S1})] \cdot R(-P_S) M_P \cdot S_{in} \quad (1)$$

where S_{in} and S_{out} are the Stokes vectors of the beam before entering the PSG and after passing through the PSG, respectively; As and P_S are the transparent-axis azimuths of the polarizers A and P, respectively; C_{S1} and C_{S2} correspond to the initial fast-axis azimuths of compensators C_1 and C_2, respectively; δ_1 and δ_2 are the phase delays of the compensators C_1 and C_2, respectively; Ma, M_P, M_{C2}, M_{C1}, and M_L are the Mueller matrices of the polarizer A, polarizer P, compensator C_2, compensator C_1, and sample to be measured, respectively; and $R(\theta)$ represents for the Mueller matrix when the element rotates θ relative to the incident plane. Without considering the polarizer rotation, light leakage, and stress birefringence, the following matrices are obtained [18] (2)–(7), as shown at the bottom of the page.

An experiment was conducted using a Mueller ellipsometry measurement system with double-rotating compensators. The measurement system is a continuous rotating polarization system. The compensators C_1 and C_2 rotate synchronously at an angular velocity ratio of 1:5. The stepper motor drives the wave plate C_1 to rotate with a fixed step size. Each position is a polarization modulation state, m is the number of polarization states generated in a period, and n is the number of detection states corresponding to each polarization modulation position. The number of light intensity measurements in an optical period is m*n. To obtain 16 elements of the Mueller matrix of samples, at least 16 sets of light intensity data must be measured, where $m^*n \geq 16$.

The first element of the Stokes vector represents the light intensity. The polarization states of the PSG and PSA are denoted by G and W, respectively. The Stokes vector of the incident light is expressed as follows: $S_{in} = (1, 0, 0, 0)^T$. The outgoing light intensity is only related to the first-row element G_i of $R(-C_{S1})M_{C1}(\delta_1)R(C_{S1})R(-P_S)M_P$ and the first-row element W_j of $R(-C_{S2})M_{C2}(\delta_2)R(C_{S2})R(A_s)M_A$. Thus, the output light intensity detected by the light intensity detector at time t, which is denoted by I(t), is expressed as follows [1]:

$$I(t) = W_1 M_L G_1 \tag{8}$$

where

$$G_i = \begin{bmatrix} 1 \\ \cos^2\frac{\delta_1}{2}\cos(2P_S) + \sin^2\frac{\delta_1}{2}\cos(4C_{S1} - 2P_S) \\ \cos^2\frac{\delta_1}{2}\cos(2P_S) + \sin^2\frac{\delta_1}{2}\sin(4C_{S1} - 2P_S) \\ \sin(2C_{S1} - 2P_S)\sin\delta_1 \end{bmatrix} \tag{9}$$

$$M_A = M_P = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} M_{C_i}(\delta_i) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\delta_i & \sin\delta_i \\ 0 & 0 & -\sin\delta_i & \cos\delta_i \end{bmatrix} \quad (i = 1, 2) \tag{2}$$

$$R(\theta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\theta & \sin 2\theta & 0 \\ 0 & -\sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (\theta = P_S, C_1, C_2, A_S) \tag{3}$$

$$R(-P_S)M_P = \frac{1}{2}\begin{pmatrix} 1 & \cos 2P_S & \sin 2P_S & 0 \\ \cos 2P_S & \cos^2 2P_S & \sin 2P_S \cos 2P_S & 0 \\ \sin 2P_S & \sin 2P_S \cos 2P_S & \sin^2 2P_S & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{4}$$

$$M_A R(A_S) = \frac{1}{2}\begin{pmatrix} 1 & \cos 2A_S & \sin 2A_S & 0 \\ \cos 2A_S & \cos^2 2A_S & \sin 2A_S \cos 2A_S & 0 \\ \sin 2A_S & \sin 2A_S \cos 2A_S & \sin^2 2A_S & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{5}$$

$$\begin{aligned} M_{C_1} &= R(-C_1) M_{c_1}(\Delta_1) R(C_1) \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1-(1-\cos\Delta_1)\sin^2 2C_1 & (1-\cos\Delta_1)\sin 2C_1 \cos 2C_1 & -\sin\Delta_1 \sin 2C_1 \\ 0 & (1-\cos\Delta_1)\sin 2C_1 \cos 2C_1 & 1-(1-\cos\Delta_1)\cos^2 2C_1 & \cos\Delta_1 \cos 2C_1 \\ 0 & \sin\Delta_1 \sin 2C_1 & -\sin\Delta_1 \cos 2C_1 & \cos\Delta_1 \end{pmatrix} \end{aligned} \tag{6}$$

$$\begin{aligned} M_{C_1} &= R(-C_1) M_{c_1}(\Delta_1) R(C_1) \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1-(1-\cos\Delta_1)\sin^2 2C_1 & (1-\cos\Delta_1)\sin 2C_1 \cos 2C_1 & -\sin\Delta_1 \sin 2C_1 \\ 0 & (1-\cos\Delta_1)\sin 2C_1 \cos 2C_1 & 1-(1-\cos\Delta_1)\cos^2 2C_1 & \cos\Delta_1 \cos 2C_1 \\ 0 & \sin\Delta_1 \sin 2C_1 & -\sin\Delta_1 \cos 2C_1 & \cos\Delta_1 \end{pmatrix} \end{aligned} \tag{7}$$

$$W_i = \begin{bmatrix} 1 \\ \cos^2\frac{\delta_2}{2}\cos(2A_S)+\sin^2\frac{\delta_2}{2}\cos(4C_{S2}-2A_S) \\ \cos^2\frac{\delta_2}{2}\cos(2A_S)+\sin^2\frac{\delta_2}{2}\sin(4C_{S2}-2A_S) \\ \sin(2C_{S2}-2A_S)\sin\delta_2 \end{bmatrix}^{\mathrm{T}} \quad (10)$$

After m*n datasets of light intensity are collected, the theoretical relationships between the light intensity matrix and the polarization modulation matrix, polarization detection modulation matrix, and sample Mueller matrix are determined.

$$\begin{aligned} I_{n*m} &= W_{n*4}M_L G_{4*m} \\ &= \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_n \end{bmatrix} \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & m_{44} \end{bmatrix} \begin{bmatrix} G_1 & G_2 & \cdots & G_n \end{bmatrix} \end{aligned} \quad (11)$$

where I_{m*n} is the light intensity matrix of all the polarization states. By performing the vectorization operation on Equation (10), the intensity matrix I_{m*n} is expanded into a column vector with m*n rows. The matrix M_L is expanded into a column vector with 16 rows.

$$\begin{aligned} \begin{bmatrix} I_{11} \\ I_{21} \\ \vdots \\ I_{n1} \\ I_{12} \\ \vdots \\ I_{nm} \end{bmatrix} &= \left(G^T \otimes W\right) Vec\,(M) \\ &= \begin{bmatrix} g_{11}W & g_{21}W & g_{31}W & g_{41}W \\ g_{12}W & g_{22}W & g_{32}W & g_{42}W \\ \vdots & \vdots & \vdots & \vdots \\ g_{1m}W & g_{2m}W & g_{3m}W & g_{4m}W \end{bmatrix} \begin{bmatrix} m_{11} \\ m_{21} \\ m_{31} \\ m_{41} \\ m_{12} \\ \vdots \\ m_{44} \end{bmatrix} \end{aligned} \quad (12)$$

The 16th order matrix is the Kronecker product of the transposition of the PSG matrix G and PSA matrix W. If a complete Mueller matrix is required, at least 16 sets of light intensity data should be collected. If n = m = 4, the 16 equations in Equation (12) can be solved strictly to obtain the Mueller matrix of the sample. In general, to improve the measurement accuracy and reduce random error, the number of light intensity measurement sets should be >16. In this situation, Equation (12) is a set of overdetermined equations composed of m * n equations. To reduce the fitting error and consider the diversity and correlation of parameters, the self-adaptive differential evolution (SADE) algorithm is used to solve the equations.

III. MODELING AND SOLUTION

In an ideal case, the light intensity calculated with the Mueller ellipsometer model should be equal to the measured light intensity. However, due to the existence of error, a certain residual exists between the theoretical and measured light intensities. The solution of the sample Mueller matrix is obtained by minimizing the total residual error of the model. Solving this matrix is essentially a global optimization problem. In this study, the least squares method is used for fitting. The parameters to be fitted are 16 matrix elements of the Mueller matrix. The objective function is the residual square sum of the theoretical and measured light intensities for all the samples.

$$\begin{aligned} \chi^2 = \phi\left(u_j\right) &= \min\left[\sum_{i=1}^{m*n}\left(I_{ti}-I_{outi}\right)^2\right] \\ &= \min\left[\sum_{i=1}^{m*n} y_i^2\right] \\ &= \min f \end{aligned} \quad (13)$$

where j is the dimension of the parameter to be optimized. The objective function can be solved with an optimization algorithm according to the principle of the least squares method, and the parameters with the least square sum of residuals can be obtained. Traditional optimization algorithms mainly include the Newton method, Gauss–Newton method, and gradient descent method, which have certain limitations for the optimization of complex structures.

A. SADE

The differential evolution (DE) algorithm is an adaptive intelligent algorithm proposed by Rainer Storm and Kenneth Price in 1997. The DE algorithm has high reliability, efficiency, and robustness. It can search continuously and parallelly in the solution space. The DE algorithm has been proved to be one of the fastest evolutionary algorithms. Equation (12) is an overdetermined system of equations; therefore, determining the value of Mueller elements with a strict solution is difficult. To overcome this problem, the SADE algorithm can be used to fit the initial solution space [18]. The DE algorithm is a multiobjective optimization algorithm. It adopts the strategies of real number coding, mutation based on difference, and competitive survival. The DE algorithm comprises the same steps as the genetic algorithm does, namely mutation, crossover, and selection. In the DE algorithm, parent individuals are randomly extracted from the initial population to mutate them and cross them with each other for forming a new vector individual. The greedy strategy is then used to select the better result between the parent individual vector and the new vector individual as the offspring. The SADE algorithm is used to obtain a higher convergence speed and accuracy than those of the DE algorithm. The flowchart of the SADE algorithm is displayed in FIGURE 2.

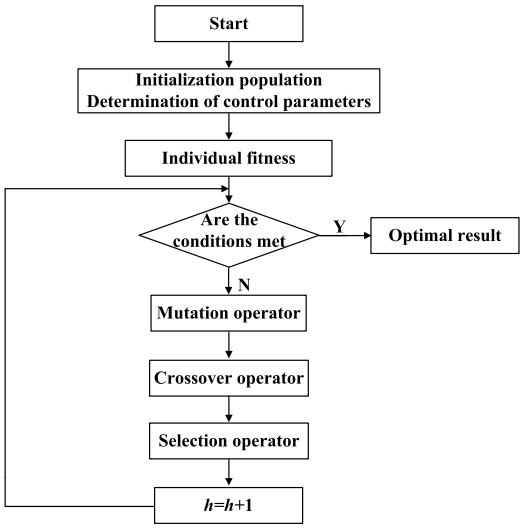

FIGURE 2. Flowchart of the SADE algorithm.

B. POPULATION INITIALIZATION

In order to make the measurement accuracy of the algorithm, we first measure the thickness of the film roughly to get the thickness interval and the initial parameter interval, and then use the SADE to explore the parameters that can make the objective function optimal in the initial parameter interval.

The number of individuals in the initial solution NP is determined according to the parameters to be fitted. For a dimension D, the individuals in the initial population are expressed as follows:

$$x_i^k = [x_{i,1}^k, x_{i,2}^k, x_{i,3}^k, \ldots, x_{i,D}^k], \quad i = 1, 2, \ldots, NP \tag{14}$$

where x_i^k is the individual of the kth iteration and 0 is the initial population. In the solution space, each individual randomly generates a set of initial vectors as the parent as follows:

$$x_{i,j}^0 = L_{j-\min} + rand_{i,j}(0, 1)\left(L_{j-\max} - L_{j-\min}\right) \quad j = 1, 2, \ldots, D \tag{15}$$

The solution space of the DE algorithm has certain boundary conditions and represents the upper and lower limits of the solution space.

C. MUTATION OPERATOR

The DE algorithm is based on the difference vector between the parent individuals for mutation operation. For the individual x_i^k in the k-generation population, three individuals except $x_{r_1}^k$, $x_{r_2}^k$, and $x_{r_3}^k$ are randomly selected. The difference vector between the two vectors of three is weighted and added to the third individual. The mutation operator v_i^k is obtained through the mutation of the third individual. The mutation strategy can be expressed as follows:

$$v_i^k = x_{r_1}^k + F * \left(x_{r_2}^k - x_{r_3}^k\right) \tag{16}$$

where $\Delta_{r_2,r_3} = x_{r_2}^k - x_{r_3}^k$ is the difference vector and F is the given scaling factor that can control the influence of the difference vector on the target individual in the mutation process. In general, F is between 0 and 2 and is usually taken as 0.5. However, to ensure the mutation of the optimal individual and improve the superiority of the mutation, F is adjusted adaptively. Suppose that the order of the three randomly selected mutation operators from the best to the worst is $x_{r_1}^k$, $x_{r_2}^k$, and $x_{r_3}^k$. Moreover, assume that the fitness values corresponding to the aforementioned operators are f_{r1}, f_{r2}, and f_{r3}, respectively. The value of F is adjusted according to the adaptive degree of two individuals corresponding to the difference vector as follows:

$$F_i = F_l + (F_u - F_l)\frac{f_{r_2} - f_{r_1}}{f_{r_3} - f_{r_1}} \tag{17}$$

The values of F_l and F_u are 0.1 and 0.9, respectively.

D. CROSSOVER OPERATOR

To increase the diversity of the population, the crossover operator is introduced. In the kth iteration, each parent is crossed with its corresponding mutation vector. In contrast to the genetic algorithm, the crossover object of the DE algorithm is the dimension of the individual. An intermediate variable or a parent is selected as the offspring according to the adaptive crossover probability.

$$h_{i,j}^k = \begin{cases} v_{i,j}^k, & rand\ (0, 1) \le P_{cr} \\ x_{i,j}^k, & else \end{cases} \tag{18}$$

where rand (0, 1) is a random number. The parameter P_{CR} is the adaptive crossover probability, which is used as the threshold to control the individual selection of the offspring and ranges from 0 to 1. If P_{CR} is large, the parent is more likely to continue as the offspring than the mutation operator is, which can accelerate the convergence speed of the function. If P_{CR} is small, the mutation operator $v_{i,j}^k$ is more likely to be the offspring than the parent is, which can enhance the species diversity. However, large or small single P_{CR} is inconducive to the iteration of the function. A suitable crossover probability should be selected for the mutated individual and next-generation mutation. In this manner, the convergence speed can be improved and the global optimal solution can be found. The adaptive strategy is expressed in Equation 3-7.

$$P_{cr_i} = \begin{cases} P_{cr_1} + \left(P_{cr_u} - P_{cr_1}\right)\dfrac{f_i - f_{\min}}{f_{\max} - f_{\min}}, & f_i \ge \bar{f} \\ P_{cr_1}, & f_i < \bar{f} \end{cases} \tag{19}$$

where f_i is the individual fitness of x_i; $\bar{f}$ is the average fitness of the current population; f_{min} and f_{max} are the fitness of the worst and best individuals, respectively; and P_{cr_1} and P_{cr_u} are the lower and upper bounds of the population adaptive probability, respectively (usually 0.1 and 0.6, respectively).

E. SELECTION OPERATOR

According to the principle of the greedy algorithm, the DE algorithm compares the target individual with the crossed

individual for each generation and selects the better individual as the offspring to enter the next iteration according to the adaptive function f.

$$x_i^{k+1} = \begin{cases} h_i^k, & f\left(h_i^k\right) < f\left(x_i^k\right) \\ x_i^k, & else \end{cases} \tag{20}$$

The individual selected according to Equation (20) must be better than the parent individual, and this individual eventually converges to an optimal solution.

IV. EXPERIMENT AND DATA ANALYSIS

A. DATA ACQUISITION AND PREPROCESSING

To measure the optical constants and thicknesses of standard film samples with the SADE algorithm, the films should meet the requirements of the ideal structure model before the simulation calculation. According to the optical properties of the sample film and substrate, a 4*4-order Mueller matrix of the optical constants, sample thickness, and incident angle is established using transfer matrix method. For isotropic nanofilms, the elliptically polarized angle (ψ) and phase difference (Δ) are used to describe the multilayered matrix as follows:

$$M_L = \begin{bmatrix} 1 & -N & 0 & 0 \\ -N & 1 & 0 & 0 \\ 0 & 0 & C & S \\ 0 & 0 & -S & C \end{bmatrix} \tag{21}$$

where

$$N = \cos 2\psi \tag{22}$$

$$S = \sin 2\psi \sin \Delta \tag{23}$$

$$C = \sin 2\psi \cos \Delta \tag{24}$$

In general, the aforementioned matrix meets the following requirement:

$$N^2 + S^2 + C^2 = 1 \tag{25}$$

B. SIMULATION CALCULATION

According to the equations in Sections II and III, the SADE characterization method was executed. SiO_2/Si standard samples with nominal thicknesses of 100.4 and 121.56 nm were measured in one optical period (π) in the wavelength range of 500–900 nm and at a light incidence angle of 65°. The step angle of the stepper motor was 1.8°. Two synchronous and continuous rotating wave plates were rotated at angular velocities of 400 and 80 r/min, with the initial point being the starting time point. The emitted light intensity was measured with the fixed step sizes N1 and N2 at the aforementioned angular velocities, respectively. A total of 81 sets of light intensity data were collected, and the number of samples n(m*n) was 81. TABLE 1 presents the azimuths of the light transmission axes of the polarizers P and A as well as the initial azimuth angles and phase delays of the rotating wave plates C_1 and C_2.

The selected initial value can directly affect the iteration speed and calculation results. The initial population of the SADE algorithm is usually generated randomly within the range of the parameter solution; however, if the initial solution deviates considerably from the accurate value, the number of calculations increases and the initial population is determined by overdetermined equations.

In practical applications, for the standard sample with an ideal isotropic film thickness, 16 elements in the Mueller matrix have specific physical significance. Among these elements, m_{11} only contains light intensity information. To directly represent the polarization state of the beam, the Stokes vector of the beam is normalized. The elements m_{13}, m_{14}, m_{23}, m_{24}, m_{31}, m_{32}, m_{41}, and m_{42} are 0 by default, which reduces the parameters to be fitted and thus the computational complexity.

The SADE algorithm is used to begin iterations for the initial population listed in Table 2. The period of sample collection is 0.75 s. The iterations are stopped when the

TABLE 1. Parameter values of the system components.

Element	Parameter	Calibration value (°)
Polarizer P	Azimuth A_S	45
Wave-plate C_1	Initial Azimuth c_1	0
Wave-plate C_1	Phase Delay Δ_1	90
Wave-plate C_2	Initial Azimuth c_2	0
Wave-plate C_2	Phase Delay Δ_2	90
Analyzers A	Azimuth P_S	-45

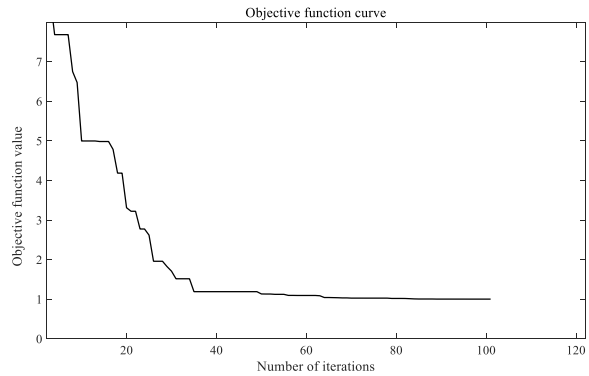

FIGURE 3. Iterative curve of the 100.4-nm-thick SiO_2/Si standard sample.

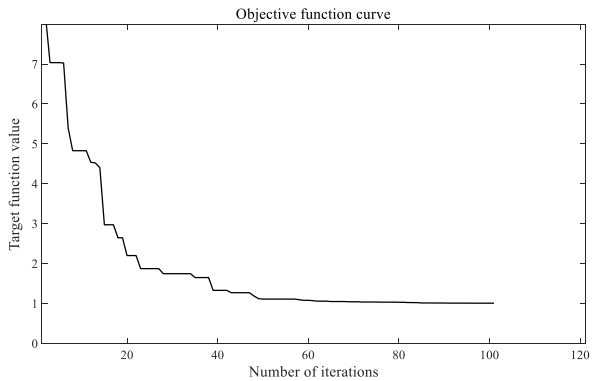

FIGURE 4. Iterative curve of the 121.56-nm-thick SiO_2/Si standard sample.

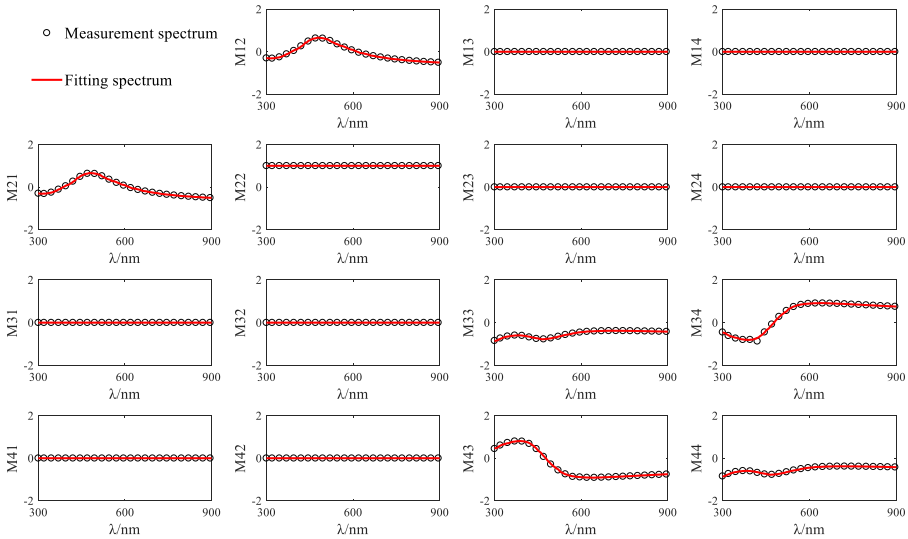

FIGURE 5. Mueller spectrum of the 100.4-nm-thick SiO_2/Si standard sample.

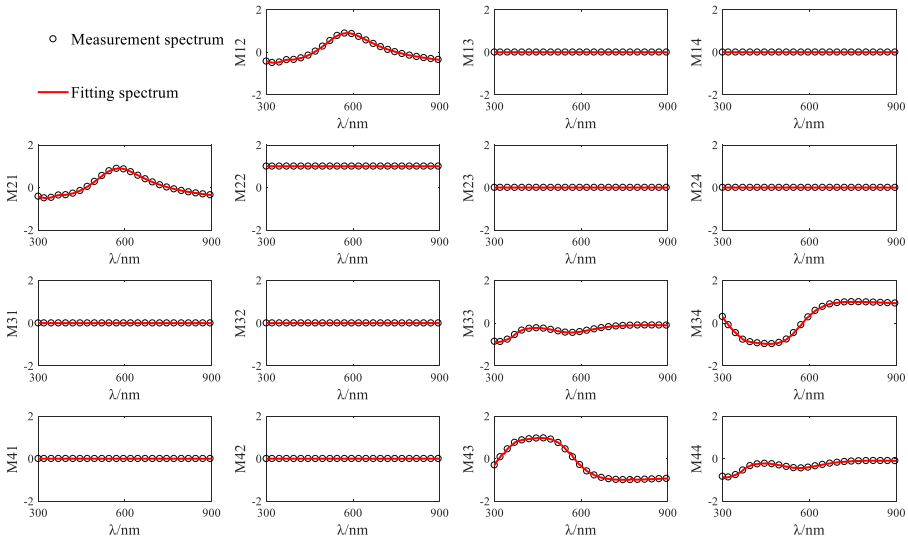

FIGURE 6. Mueller spectrum of the 121.56-nm-thick SiO_2/Si standard sample.

convergence value is <1 or the maximum number of iterations is 100. The iteration curve is displayed in figure 3.

The iterative curves of the objective function indicate that when the 100.4-nm-thick and 121.56-nm-thick standard samples are iterated 65 and 80 times, respectively, the sums of the squares of the residuals converge to the minimum values of 1.24 and 1.02, respectively. Moreover, for the 100.4-nm-thick and 121.56-m-thick standard samples, the mutation operator f is 0.6 and 0.5, respectively, and the crossover probability P_{CR} is 0.1 and 0.3, respectively. Thus, the SADE algorithm has a high convergence speed for calculating the exact value of objective function in a short time. Moreover, it can prevent the calculation result from falling into the local optimal solution and can find the global optimal solution by constantly adjusting the step size.

Comparisons between the spectra of the sample Mueller matrix obtained through fitting and those measured with the Muller matrix ellipsometry measurement system (dual-rotating compensator Mueller matrix ellipsometer) are displayed in FIGURE 6 and 7. The fitting errors for the aforementioned two methods are 36.8 and 32.4, respectively. The depolarization error is one of the main error sources in the calculation of the Mueller matrix. The depolarization index (DI; $0 \leq DI \leq 1$) of the Mueller spectrum is 0.9316 and 0.932 at wavelengths of 300 and 900 nm, respectively. The order of magnitude of the DI is equivalent to that of the calculation error; thus, the depolarization effect can be ignored.

TABLE 2. Calculated parameters for the two standard samples.

Nominal thickness / nm	Thickness calculation value / nm	relative error
100.41	101.25	0.84%
121.56	120.53	0.85%

TABLE 3. Results of repeated experiments.

Nominal thickness / nm	Method	Average thickness calculation value / nm	Max relative error	Min relative error	Average relative error	Average Time/s
100.40nm	SADE	101.25	0.97%	0.46%	0.84%	1.41
	Newton	96.04	5.11%	3.94%	4.31%	16.4
	GN	97.85	4.39%	2.27%	2.53%	17.4
	GD	93.21	11.45%	2.18%	8.07%	25.3
121.56nm	SADE	120.53	0.91%	0.57%	0.84%	1.32
	Newton	113.22	8.16%	5.71%	6.86%	29.3
	GN	117.36	7.13%	3.22%	3.46%	36.8
	GD	109.18	11.31%	4.35%	10.18%	42.7

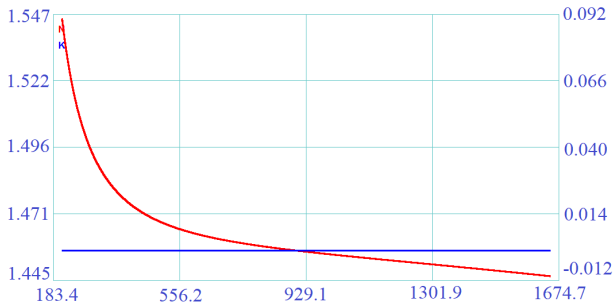

FIGURE 7. Variation in the refractive index of the 100.40-nm-thick SiO2/Si standard sample with the wavelength of the incident light.

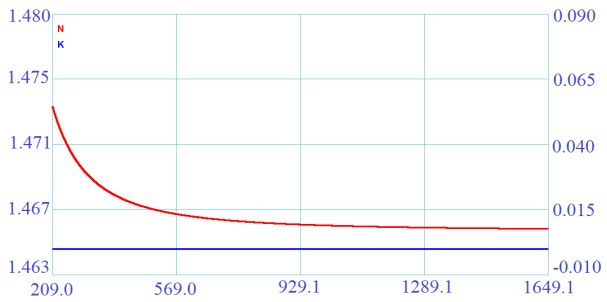

FIGURE 8. Variation in the refractive index of the 121.56-nm-thick SiO2/Si standard sample with the wavelength of the incident light.

For isotropic standard samples, the elements containing the optical constants and structural information are M_{12}, M_{21}, M_{33}, M_{34}, M_{43} and M_{44}. On the basis of the spectral curves of the aforementioned elements, the polarization angle (ψ) and phase difference (Δ) of the sample are calculated using the inversion method. However, to avoid the error of the second inversion, the transmission matrix formula (Equation 21) is directly introduced into the light intensity formula for obtaining an iterative solution. The film thickness (d), film refractive index (n), and extinction coefficient (k) can be obtained using the associated ellipsoid equation. When SiO_2/Si thin films are transparent, the extinction coefficient k is 0. Therefore, the ellipsometer can be used to obtain the films and refractive index after measuring the obtained ellipsometry parameters.

FIGURE 7 and FIGURE 8 respectively show the change of refractive index of 100.40 nm and 121.56 nm standard samples with the wavelength of incident light measured by Wuhan Yiguang ME-L ellipsometer (Wuhan Yiguang ME-L ellipsometer, Wuhan eoptics Technology Company, Wuhan, China). The refractive index at the wavelength of 600 nm was selected to calculate the film thickness value of the sample and conduct analysis.

The thickness values obtained after the National Institute of Standards and Technology (NIST) calibration of the very large-scale integration samples are 100.41 $\pm$ 0.4 and 121.56 $\pm$ 0.4 nm. The simulation results of the SADE algorithm are within the allowable range of calibration error, and the calculation error for the thickness is $<1\%$. Moreover, the calculation error of the SADE algorithm is almost equal to that of the Mueller element. This finding verifies the accuracy of the film thickness results obtained using the SADE algorithm. Because the SADE algorithm adjusts the step size of the next iteration according to the fitness of each parent population, it can quickly find an appropriate search direction for the target solution. The SADE algorithm can not only shorten the solution time but also prevent the solution from falling into periodic local optima. It can also be used to accurately calculate the film thickness. The SADE algorithm is suitable for complex optimization problems with multiobjective solutions.

C. COMPARATIVE EXPERIMENTS

Because the model established in this paper is an unconstrained problem, to solve the problem we need to find the best parameter vector in the limited parameter interval after many times of optimization. We found the most commonly used method to carry out the comparison experiment.

In order to compare the computational effect of SADE algorithm and traditional optimization algorithm, we use Newton's method, Gauss Newton (GN) method and Gradient Descent (GD) method to compare with SADE algorithm. We use Wuhan Yiguang ME-L ellipsometer to collect the light intensity values of standard samples, and use MATLAB software (MATLAB, MathWorks, Natick, MA, USA) to compile the above algorithm on the computer with AMD Ryzen 5 4600H processor, and calculate the matrix parameters of 100.4 nm and 121.56 nm SiO_2/Si standard samples. Then the matrix parameters fitted by the algorithm are input into the refractive index calculation software of Wuhan Yiguang ME-L ellipsometer. After obtaining the

TABLE 4. Results of repeated experiments.

Nominal thickness / nm	Method	Average thickness calculation value / nm	Max relative error	Min relative error	Average relative error	Average Time/s
50.00nm	SADE	49.69	0.86%	0.54%	0.63%	1.37
	LM	49.61	0.81%	0.66%	0.78%	1.01
	IPSO-NN	49.45	1.30%	0.94%	1.10%	4.69
	MBO	49.27	4.46%	0.59%	1.46%	7.76
	EHO	49.32	3.37%	0.44%	1.37%	8.12
100.40nm	SADE	101.25	0.97%	0.46%	0.84%	1.41
	LM	99.45	1.16%	0.77%	0.95%	1.22
	IPSO-NN	98.88	1.84%	1.21%	1.51%	4.52
	MBO	97.98	4.79%	0.32%	3.37%	8.34
	EHO	99.83	2.07%	0.32%	1.55%	12.32
121.56nm	SADE	120.53	0.91%	0.57%	0.84%	1.32
	LM	120.10	1.33%	1.01%	1.20%	1.13
	IPSO-NN	119.70	1.92%	1.18%	1.53%	5.02
	MBO	117.07	4.15%	0.61%	3.69%	11.19
	EHO	119.17	3.63%	0.46%	2.01%	8.97
997.70nm	SADE	982.23	1.83%	1.41%	1.55%	1.57
	LM	980.69	1.97%	1.66%	1.72%	1.36
	IPSO-NN	971.96	2.67%	2.48%	2.58%	3.96
	MBO	955.90	5.59%	2.30%	4.19%	9.15
	EHO	959.29	4.11%	1.45%	3.85%	10.29

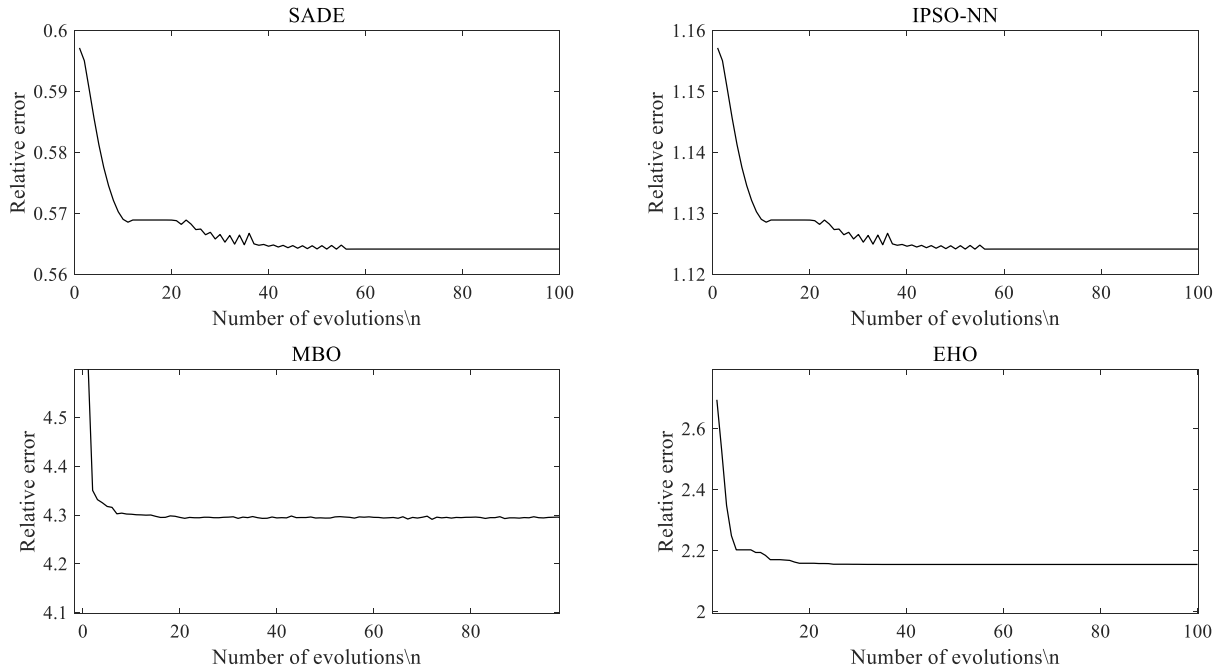

FIGURE 9. Iteration results of different algorithms (50.00-nm-thick SiO2/Si standard sample).

curve of refractive index changing with the wavelength of incident light, the experimental parameters of film thickness with wavelength of 600 nm are selected. Repeat the above operation to get the average value of 10 times fitting and the measurement results corresponding to the average value, as shown in Table 3.

It can be seen from Table 3 that Newton's method, GN and GD all are slow in solving the problems, and the calculation results are not accurate. This is because we need to solve so many parameters, and the sample size is large, our algorithm is easy to fall into local optimum. SADE algorithm in this paper is a kind of evolutionary algorithm. It has advantages for solving the problem of many parameters and large sample size. It can effectively avoid the extreme optimal value and find the optimal value.

In order to compare the influence of SADE algorithm with other metaheuristic algorithms on the calibration accuracy, we use the latest algorithm for characterization of thin film parameters [19] and the popular meta-heuristic algorithms Monarch butterfly optimization (MBO) [20], Elephant herding optimization (EHO) [21] in recent years. We also use Wuhan Yiguang ME-L ellipsometer to collect the

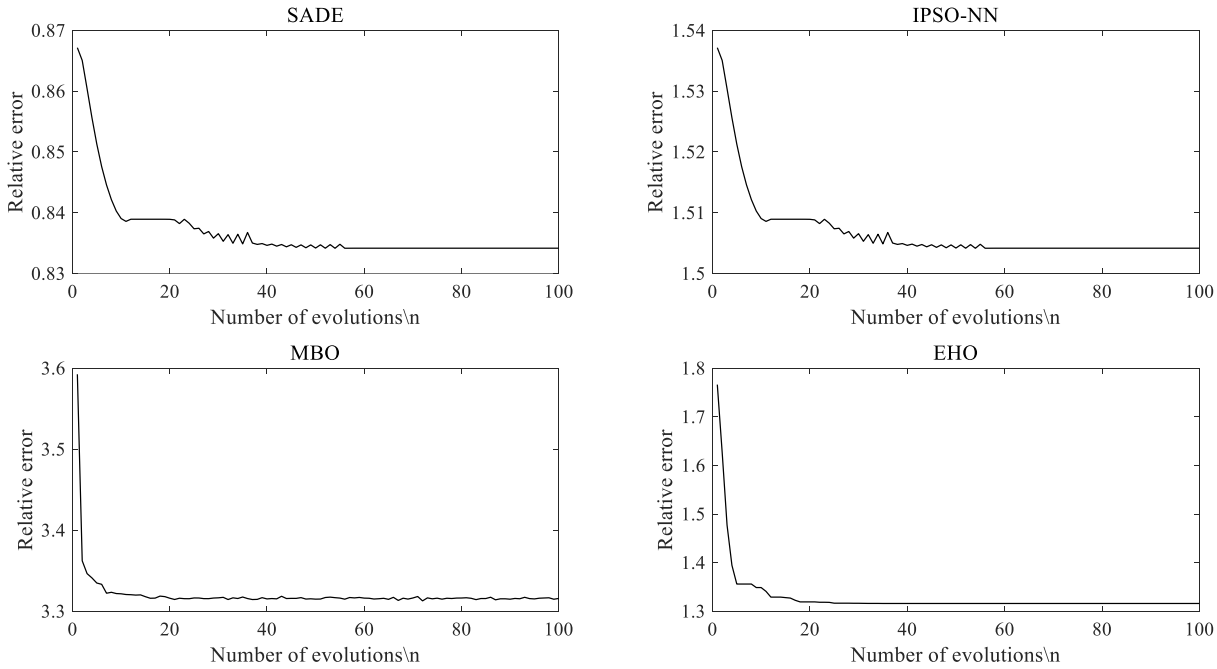

FIGURE 10. Iteration results of different algorithms (100.4-nm-thick SiO2/Si standard sample).

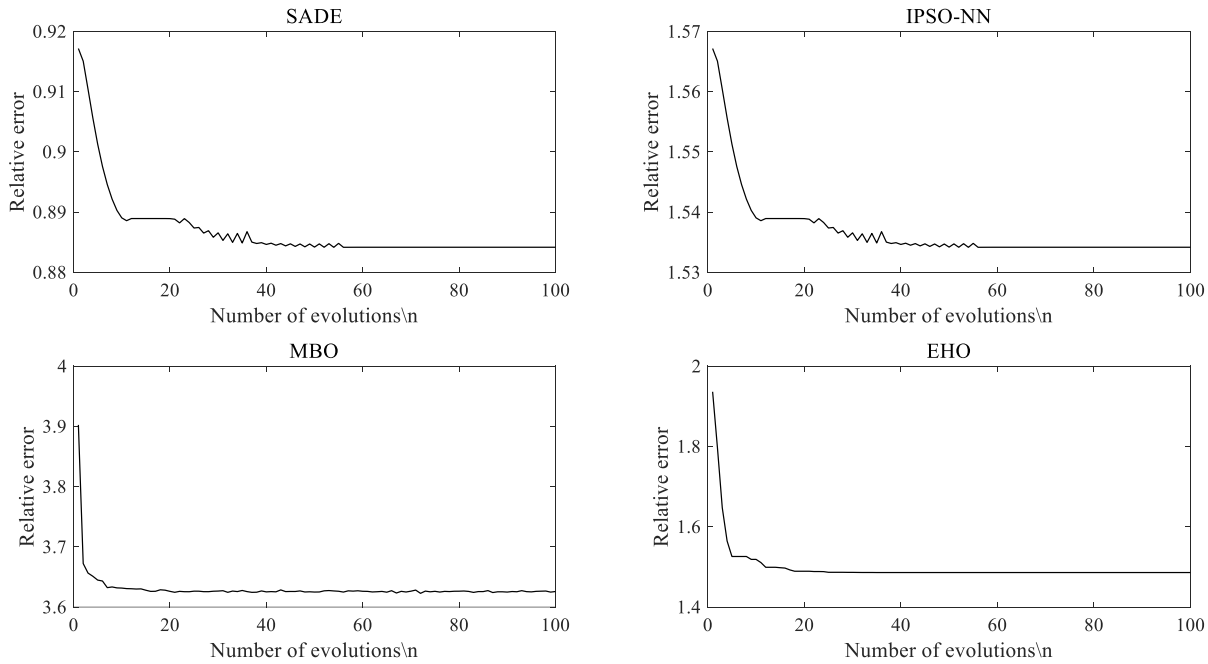

FIGURE 11. Iteration results of different algorithms (121.56-nm-thick SiO2/Si standard sample).

light intensity values of standard samples, and use MATLAB software to compile the above algorithm on the computer with AMD Ryzen 5 4600H processor. In the contrast experiment, MBO's elitism parameter is 2, max step size is 1, period is 1.2, and the number of iterations is 100; IPSO-NN's learning factor is 2, initial weight is 0.8, decreasing coefficient is 10^{-4}, NN's learning parameters are 0.5, and the number of iterations is 100. The EHO's elitism parameter is 2, the number of search agents is 50, and the number of iterations is 100.In order to make the results more reliable, we calculated the matrix parameters of SiO_2/Si standard samples at 50.00 nm, 100.4 nm, 121.56 nm and 997.70 nm. After obtaining the curve of refractive index changing with the wavelength of incident light, we also selected the experimental parameters of film thickness at 600 nm. In addition, we also use the supporting software packages

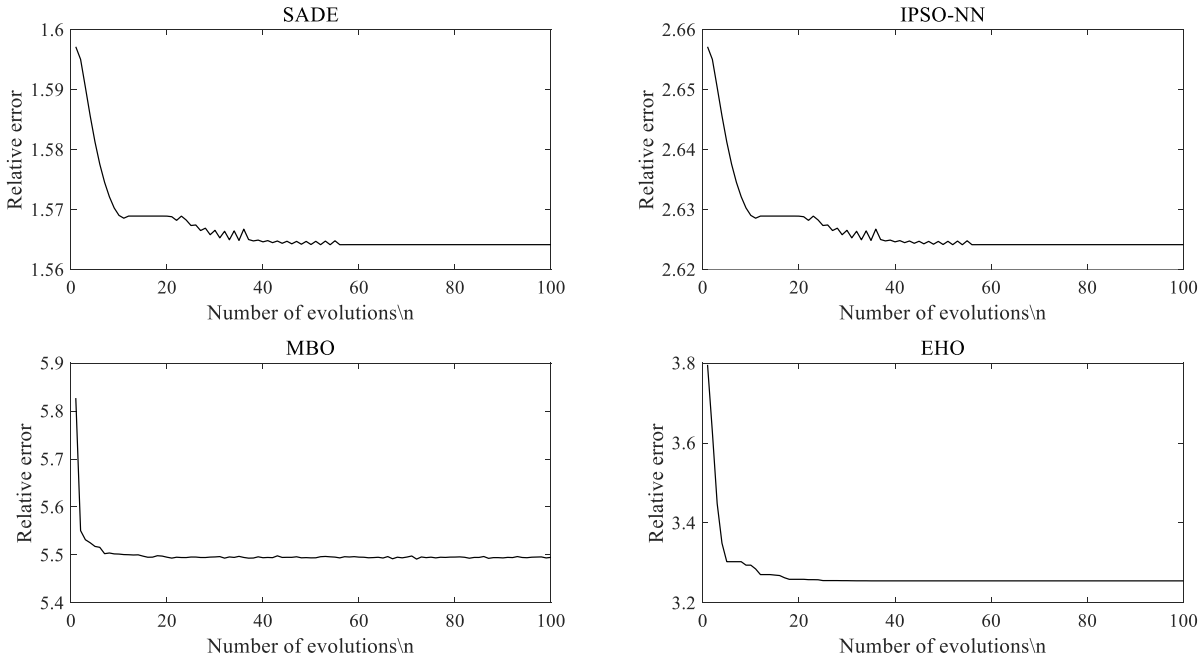

FIGURE 12. Iteration results of different algorithms (997.70-nm-thick SiO2/Si standard sample).

of CompleteEASE (CompleteEASE, JA woolam Company, Lincoln, NE, USA), which based on Levenberg-Marquardt (LM) method for comparison with above methods. Repeat the above operation to get the average value of 10 times fitting and the measurement results corresponding to the average value, as shown in table 4

As can be seen from Table 4, FIGURE 9, 10, 11 and 12, when solving the problem with multiple parameters and large sample data, the meta-heuristic algorithm can indeed effectively avoid the extreme value and find the global optimal value. MBO and EHO are particularly fast, and the region near the optimal solution can be found within 20 generations. However, they are not as accurate as SADE, LM and PAO-NN in judging global and local optimality in the later stage of program running. This is because the maximum search step size of MBO and EHO has an obvious effect on the results. When the step size is small, the global search ability of the algorithm is poor, and the local search ability is strong, and the accuracy is higher. Conversely, the global search ability is strong, and the local search ability is poor, and the accuracy is lower but more stable. SADE, LM and PSO-NN are the algorithms that have been debug-tested and verified for many times in this project. However, the calculation method of PSO-NN is more complex than SADE. In the case of the same number of iterations, PSO-NN's results are not as stable as SADE. Although LM method, have been proven to be stable and adaptive to other engineering applications with the ability to accurately determine both the thickness and complex optical constants across a wide variety of materials. However, SADE has the ability to distinguish different features from big data and can adapt to the features obtained from multi-parameter data by Muller method, so that SADE's calculations are more accurate.

V. CONCLUSION

This paper proposes the use of the SADE algorithm (The number of iterations is 50, the square of residual error is 1.24, the minimum value of convergence is 1.02, the mutation operator F is 0.6, and the crossover probability Pcr is 0.1.) to solve the Muller matrix of the thickness of a standard nanofilm for calculating multiple film thickness parameters and optical constants.

First, considering the diversity of the parameters to be fitted, the SADE algorithm is used to solve the overdetermined equations.

Second, the simulation results are compared with NIST calibration results. The relative errors of the two samples are $<1\%$. This result indicates that the SADE algorithm is convenient for calculations and has high accuracy in characterizing isotropic nanofilms.

Third, this paper uses the adaptive differential evolution algorithm to solve the model. Through experiments, compared with the traditional algorithm and the existing method, it has a smaller error value in the calculation of the data collected by Muller ellipsometer.

Fourth, in repeated experiments, we find that other meta heuristic algorithms can achieve smaller relative error, which we think is because the random algorithm finds a more suitable initial value. However, the calculation results of MBO and EHO are not stable and can not adapt to long-term engineering application. Moreover, Sade has shorter computing time and is more suitable for poor hardware conditions.

Finally, in this paper, the accuracy of iteration depends greatly on the selection of initial value. In order to select appropriate initial value, the initial value of iteration should be set according to historical experience when detecting conventional samples to be tested, while rough measurement

should be carried out first to get the approximate range of parameter value when detecting unconventional samples to be tested.

REFERENCES

[1] L. R. Dahal, J. Li, J. A. Stoke, Z. Huang, A. Shan, A. S. Ferlauto, C. R. Wronski, R. W. Collins, and N. J. Podraza, "Applications of real-time and mapping spectroscopic ellipsometry for process development and optimization in hydrogenated silicon thin-film photovoltaics technology," *Sol. Energy Mater. Sol. Cells*, vol. 129, pp. 32–56, Oct. 2014, doi: 10.1016/j.solmat.2014.01.028.

[2] C. Chen, I. An, G. M. Ferreira, N. J. Podraza, J. A. Zapien, and R. W. Collins, "Multichannel Mueller matrix ellipsometer based on the dual rotating compensator principle," *Thin Solid Films*, vols. 455–456, pp. 14–23, May 2004, doi: 10.1016/j.tsf.2003.11.191.

[3] S. Liu, X. Chen, and C. Zhang, "Development of a broadband Mueller matrix ellipsometer as a powerful tool for nanostructure metrology," *Thin Solid Films*, vol. 584, pp. 176–185, Jun. 2015, doi: 10.1016/j.tsf.2015.02.006.

[4] E. A. Irene and D. W. Dong, "Ellipsometry measurements of polycrystalline silicon films," *J. Electrochem. Soc.*, vol. 129, no. 6, pp. 1378–1386, 1982, doi: 10.1149/1.2124148.

[5] Z. Chen, R. Meng, Y. Zhu, and H. Ma, "A collinear reflection Mueller matrix microscope for backscattering Mueller matrix imaging," *Opt. Lasers Eng.*, vol. 129, Jun. 2020, Art. no. 106055, doi: 10.1016/j.optlaseng.2020.106055.

[6] J.-H. Kim and M. Swaminathan, "Modeling of irregular shaped power distribution planes using transmission matrix method," *IEEE Trans. Adv. Packag.*, vol. 24, no. 3, pp. 334–346, 2001, doi: 10.1109/6040.938301.

[7] H. B. Kim, J. H. Son, C. N. Whang, and K. H. Chae, "Ellipsometric spectroscopy study of air ion-beam mixed $SiO_2/Si/SiO_2$ layers," *Nucl. Instrum. Methods Phys. Res. B, Beam Interact. Mater. At.*, vol. 216, pp. 367–371, Feb. 2004, doi: 10.1016/j.nimb.2003.11.062.

[8] J. Li, X. Sun, Q. Kang, S. Li, and Y. Yin, "Polarization detection accuracy analysis of spectropolarimeter," *Infr. Laser Eng.*, vol. 47, no. 1, 2018, Art. no. 0123002, doi: 10.3788/IRLA201847.0123002.

[9] D. E. Aspnes, "Spectroscopic ellipsometry—Past, present, and future," *Thin Solid Films*, vol. 571, pp. 334–344, Nov. 2014, doi: 10.1016/j.tsf.2014.03.056.

[10] J. Li, B. Ramanujam, and R. W. Collins, "Dual rotating compensator ellipsometry: Theory and simulations," *Thin Solid Films*, vol. 519, no. 9, pp. 2725–2729, Feb. 2011, doi: 10.1016/j.tsf.2010.11.075.

[11] G. Yuqing, T. Dongmei, F. Yunxia, S. Jiayuan, H. Zhiguo, Z. Bo, K. Ming, C. Chengming, and L. Lihua, "Study on LM algorithm in Mueller's ellipsometry calibration method," *Infr. Laser Eng.*, vol. 49, no. 8, pp. 168–176, 2020, doi: 10.3788/IRLA20200204.

[12] R. Battiti, "First- and second-order methods for learning: Between steepest descent and Newton's method," *Neural Comput.*, vol. 4, no. 2, pp. 141–166, Mar. 1992, doi: 10.1162/neco.1992.4.2.141.

[13] R. W. M. Wedderburn, "Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method," *Biometrika*, vol. 61, no. 3, pp. 439–447, 1974, doi: 10.2307/2334725.

[14] L. M. S. Aas, P. G. Ellingsen, B. E. Fladmark, P. A. Letnes, and M. Kildemo, "Overdetermined broadband spectroscopic Mueller matrix polarimeter designed by genetic algorithms," *Opt. Exp.*, vol. 21, no. 7, pp. 8753–8762, 2013, doi: 10.1364/OE.21.008753.

[15] E. Passaglia, R. R. Stronberg, J. Kruger, and M. Cohen, "Ellipsometry in the measurement of surfaces and thin films," *J. Electrochem. Soc.*, vol. 112, no. 4, p. 109C, 1965, doi: 10.1149/1.2423550.

[16] N. J. Bulitka, J. D. Leslie, and J. L. Ord, "Simulation of layer thickness measurements of quantum well structures by *in situ* single-wavelength ellipsometry," *J. Electrochem. Soc.*, vol. 139, no. 10, pp. 2895–2899, Oct. 1992, doi: 10.1149/1.2069002.

[17] R. M. A. Azzam, "Photopolarimetric measurement of the Mueller matrix by Fourier analysis of a single detected signal," *Opt. Lett.*, vol. 2, no. 6, p. 148, Jun. 1978, doi: 10.1364/OL.2.000148.

[18] F. Zhao, X. He, G. Yang, W. Ma, C. Zhang, and H. Song, "A hybrid iterated local search algorithm with adaptive perturbation mechanism by success-history based parameter adaptation for differential evolution (SHADE)," *Eng. Optim.*, vol. 52, no. 3, pp. 367–383, Mar. 2020, doi: 10.1080/0305215X.2019.1595611.

[19] L. Lihua, Z. Xinyin, W. Junjie, L. Zhiwei, L. Qiang, L. Na, X. Zhangning, G. Yuqing, and F. Yunxia, "Characterization of nanofilm parameters based on hybrid optimization algorithm," *Infr. Laser Eng.*, vol. 49, no. 2, 2020, Art. no. 213002, doi: 10.3788/IRLA202049.0213002.

[20] G.-G. Wang, S. Deb, and Z. Cui, "Monarch butterfly optimization," *Neural Comput. Appl.*, vol. 31, no. 7, pp. 1995–2014, Jul. 2019, doi: 10.1007/s00521-015-1923-y.

[21] G.-G. Wang, S. Deb, and L. D. S. Coelho, "Elephant herding optimization," in *Proc. 3rd Int. Symp. Comput. Bus. Intell. (ISCBI)*, Dec. 2015, pp. 1–5. [Online]. Available: https://www.researchgate.net/publication/282822637_Elephant_Herding_Optimization

LIHUA LEI (Member, IEEE) received the B.S. degree in measurement and control technology, and instrument and the M.S. degree in instrumental science and technology from China Jiliang University, Hangzhou, China, in 2009 and 2012, respectively, and the Ph.D. degree in physics from Tongji University, Shanghai, China, in 2015. From 2012 to 2020, his main research direction is micro and nanometrology. He has won the Gold Medal of Shanghai Excellent Invention, the Shanghai Science and Technology Progress Award three times, the Xingjian Award of AQSIQ twice, and the Science and Technology Progress Award of China Metrology and Testing Society.

YUQING GUAN was born in Suzhou, Jiangsu, China, in 1995. She received the B.S. and M.S. degrees in aerospace engineering from China Jiliang University, China, in 2021.

YANHUA ZENG graduated from China Jiliang University. She is currently working with the Shanghai Institute of Metrology and Testing Technology. Her research direction is precision instrument measurement.

LIN ZHAO was born in 1983. She received the B.S. degree in measurement and control technology, and instrument from Yanshan University, Qinhuangdao, China, in 2006, and the M.S. degree in instrumental science and technology from Tianjin University, Tianjin, China, in 2008.

She has been working as a Senior Engineer with The 13th Institute of CETC, since July 2008. As the first inventor, she has obtained one authorized invention patent and one authorized utility model patent. She has published one EI retrieval article and four core journal articles. In 2020, she won the Third Prize of Science and Technology Progress Award of Metrology Testing Society, and the Third Prize of Technological Invention Award of State Administration of Science, Technology and Industry for National Defence.

ZHANGNING XIE was born in Wenzhou, Zhejiang, China, in 1995. He received the B.S. and M.S. degrees in aerospace engineering from China Jiliang University, China, in 2020.

From 2018 to 2019, he won the National First Prize of the 14th "Zhaoyi Innovation Cup" China Graduate Student Electronic Design Competition and the Second Prize of the 16th "Huawei Cup" China Research Mathematical Modeling Competition.

ZHIGUO HAN was born in 1982. He received the B.S. degree in measurement and control technology and instrument from China Jiliang University, China, in 2005.

He has been working as a Senior Engineer with The 13th Institute of CETC, since July 2005. As the first inventor, he obtained two authorized invention patents and two software copyrights. He has published two EI retrieval articles and three core journal articles. In 2020, he was a second recipient, where he won the Third Prize of Science and Technology Progress Award of Metrology Testing Society and the Third Prize of Technological Invention Award of State Administration of Science, Technology and Industry for National Defence.

CHENGMING CAO was born in 1972. He received the bachelor's degree in applied chemistry from Shanghai JiaoTong University, Shanghai, China. He is currently the Vice President of Shanghai Institute of Measurement and Testing Technology.

YUNXIA FU was born in Henan. She is currently a Senior Engineer, an Metrological Testing Expert, and a Professor with the Shanghai Institute of Measurement and Testing Technology. She is also a member of the National Geometric Length Measurement Technical Committee and the Optical Testing Professional Committee of the Chinese Optical Society. She has been responsible for a number of scientific research projects, two of which won the Second Prize of Shanghai Science and Technology Progress Award. She is also the Vice Chairman of the Working Group of Nanometer Testing Technology Standardization and the National Technical Committee for Nanotechnology Standardization. She is also a Provincial Evaluation Member of the Nation Allegal Metrological Verification Institution, a second-level Evaluation Member of National Metrological Standard, a national metrological permit evaluation member, and an Expert of Shanghai Municipal Government Procurement Review and Evaluation and Shanghai Municipal Engineering Equipment Supervision Bidding and Tendering Expert Database. As an editorial board member, she has participated in the compilation of the national length metrology personnel training materials, and has published about 15 articles in professional journals at home and abroad in both Chinese and English.