Genetic algorithm for ellipsometric data inversion of absorbing layers

Gabriel Cormier and Roger Boudreau

École de génie, Université de Moncton, Moncton, Nouveau-Brunswick, Canada E1A 3E9

Received April 1, 1999; revised manuscript received September 22, 1999; accepted September 22, 1999

A new data reduction method is presented for single-wavelength ellipsometry. A genetic algorithm is applied to ellipsometric data to find the best fit. The sample consists of a single absorbing layer on a semi-infinite substrate. The genetic algorithm has good convergence and is applicable to many different problems, including those with different independent measurements and situations with more than two angles of incidence. Results are similar to those obtained by other inversion techniques.
[S0740-3232(00)02201-8]

OCIS codes: 120.2130, 310.6860.

1. INTRODUCTION

Ellipsometry is a well-known technique used to determine the optical properties of thin films. It is based on the principle that polarized light changes polarization state when reflected. Once a sample is irradiated with light of known polarization, wavelength, and incidence angle, information regarding the optical properties of the film and its thickness may be collected.

The ellipsometric equation is well known and is presented in many texts. If the measured sample can be treated as a bulk substrate with a perfect surface, the reflection ratio can be expressed as

$$\rho = r_p/r_s = \tan(\psi)\exp(i\Delta), \tag{1}$$

where r_p and r_s are the Fresnel reflection ratios for the components parallel (p) and perpendicular (s), respectively, to the plane of incidence. Written in polar form, the complex ratio can be expressed with the two parameters ψ and Δ (the ellipsometric angles), where $\tan\psi$ and Δ describe the amplitude ratio and the phase difference, respectively, between the p and s components. If the sample consists of a substrate with one or two more films, the measured reflection ratio can be written as

$$R_p/R_s = \tan\psi\exp(i\Delta), \tag{2}$$

where the reflection coefficients R_p and R_s are functions of the Fresnel reflection coefficients for the interfaces and the film thicknesses.[1]

For standard single-wavelength ellipsometry of an absorbing film with a known substrate, there are usually three unknowns: n, the real part of the refractive index; k, the imaginary part of the refractive index (the extinction coefficient); and d, the film thickness. Once the angles ψ and Δ have been measured, the problem is then to find an algorithm to inverse this data and determine the three unknowns, since no direct analytical solution can be found. The ellipsometric angles are measured at more than one incidence angle.

Standard inversion algorithms require a starting point, or initial value. This initial value must usually be near the correct value or the algorithm does not converge and computation time may be high, although with the exponentially rapid development of computers, the latter requirement is less of a concern.

We propose to use a genetic algorithm to inverse ellipsometric data. Genetic algorithms do not require a starting point and in general have a greater range of convergence than other inversion techniques. They are readily adaptable to many different problems and also have low computation times. They also do not require the evaluation of derivatives, instead relying on a merit function to improve performance.

2. BACKGROUND

Some of the first methods used to solve the ellipsometric inversion problem were polynomial inversion techniques. Among the first were those of McCrackin and Colson.[2] Their algorithm determined two of the parameters, n and d, when a third, k, was known. The algorithm still required an initial value, and this value had to be close to the solution. However, uncertainties in the measurements and in the method's precision rendered this algorithm impractical. Other researchers, such as Reinberg[3] and later Easwarakhanthan *et al.*,[4] used a multidimensional Newton algorithm (or a variation) to compute the film properties. Computing time was good, but this technique had some convergence problems, especially for films less than 40 nm thick.

More recently, Urban[5] proposed a new method, using an algorithm of variably damped least squares. He computed the intersection of two solution curves at different incidence angles. Two intersections are found, one for a plot of n versus d and one for d versus k. These curves intersect at the correct solution of (n, k, d). If the model is correct and the data are precise, the two solutions should be almost identical. Drolet *et al.*[6] proposed a new approach to the inversion problem. They considered the case of nonabsorbing layers ($k = 0$) and separated the problem into two steps: One step implies solving an equation that depends only on the refractive index, and

0740-3232/2000/010129-06$15.00

the second step is finding the film thickness. They reduced the refractive index equation to a fifth-degree polynomial and obtained a solution by finding its roots. Results were good and computation time was small, although problems were encountered that are related mainly to root solving and are not specific to the ellipsometric equations.

The Levenberg–Marquardt algorithm can also be used for spectroscopic ellipsometry.[7] It is a nonlinear least-squares regression method and is thus very good at handling many noisy measurements. Care is warranted if the Levenberg–Marquardt algorithm is used, since it may get caught in a local minimum.

An interesting paper was presented by Bosch *et al.*[8] They proposed modifying a downhill simplex algorithm to inverse data; their algorithm better integrated experimental data and used a better fitting procedure. In order to converge, the algorithm tries to minimize a certain function, in this case the biased estimator, recommended by Jellison[7] as a better interpretation technique. A plot of the merit function (the function to minimize) versus one of the three unknowns was made. The solution corresponds to the minimum of this plot. These curves were obtained by keeping one of the three unknowns constant and finding a value for the other two unknowns and computing the value of the merit function. Results obtained were very good even in the presence of noise.

3. GENETIC ALGORITHM

Genetic algorithms (GA's) were initially developed by Holland.[9] They are based on the mechanics of natural evolution and natural genetics and differ from usual inversion algorithms because they do not require a starting value. They use a survival-of-the-fittest scheme with a random organized search to find the best solution to a problem. They have a variety of applications[10] and are as robust as other inversion techniques. Genetic algorithms are also easily applicable to problems with many unknowns. Although they have been used in many different fields, genetic algorithms are relatively new to the field of optics. One such application can be found in Ref. 11.

The basic genetic algorithm has the following structure:

```
Create an initial population
Repeat
    Selection
    Crossover
    Mutation
Verify whether end condition is met
```

A. Initial Population

An initial population is first created, randomly, of η individuals. Each individual represents a possible solution to the problem: in this case, values of (n, k, d) that solve the ellipsometric equations. An individual is composed of chromosomes, and each chromosome is composed of genes. For this problem, individuals have only one chromosome (and thus the terms individual and chromosome are interchangeable here) and are composed of three genes, since we are solving for three variables. Individuals are usually coded as bit strings called binary-coded GA's.[10] Davis[12] has shown that real-coded GA's usually outperform GA's coded with bits, and such individuals are represented here with real numbers. The individuals that constitute the initial population are created within a specified solution interval or boundary. For example, given bounds of $2.5 < n < 3.5$, $0 < k < 0.1$, and $10 < d < 30$, two individuals might be

n	k	d
2.74	0.054	12.8
3.19	0.006	28.3

B. Selection

Once the initial population is created, each individual's performance is evaluated. The performance of an individual is a measure of how "good" this particular solution to the problem is. This measure is created by calculating the value of the objective function, the function to minimize. For ellipsometric data reduction, the objective function is the merit function χ^2,

$$\chi^2 = \frac{1}{N}\sum_{m=1}^{N}\left[\left(\frac{\Delta_m - \Delta_c}{\epsilon^{\Delta}}\right)^2 + \left(\frac{\psi_m - \psi_c}{\epsilon^{\psi}}\right)^2\right], \qquad (3)$$

where N is the number of measurements and ϵ^{Δ} and ϵ^{ψ} are the experimental errors of Δ and ψ, respectively. The subscript m indicates the measured values, and the subscript c indicates computed values. This function has interesting qualities, since it is a useful indicator of statistical significance. The fit is good if $\chi^2 \approx 1$, meaning that the errors are of the same magnitude as the measurements. Should $\chi^2 \gg 1$, the fit deviates significantly from the experimental data. Should $\chi^2 \ll 1$, the error may have been overestimated.[7]

The parallel and perpendicular reflection coefficients, R_p and R_s, respectively, can be computed for each individual in the population from the values of n, k, and d. The amplitude ratio ψ_c and the phase difference Δ_c can then be determined, and the value of χ^2 is computed and becomes the individual's fitness. Once each individual's fitness is computed, the population is ranked and each individual is given a probability of reproduction, PR.

A few ranking methods exist, but the one used was developed by Davis.[12] The individual with the best performance receives a relative weight (RW) of η^a, where a is usually chosen between 1.0 and 1.5. The second-best individual receives a relative weight of $(\eta - 1)^a$, and so on until the worst individual receives a relative weight of 1. The PR for each individual is then computed as follows:

$$\mathrm{PR}_j = \frac{\mathrm{RW}_j}{\dfrac{1}{\eta}\displaystyle\sum_{j=1}^{\eta}\mathrm{RW}_j}. \qquad (4)$$

As can be seen, the fittest individuals have a higher probability of reproduction than do those with a lesser performance. A stochastic-remainder selection procedure[10] is then used to determine the frequency of reproduction. Only the integer part of the PR is considered. For example, a PR of 2.1 means that the individual will reproduce twice, and an individual of PR 1.9 will reproduce only once. The rest of the population is made up of the sorted list of fractional parts.

C. Crossover

To generate offspring, a pair of individuals is randomly selected according to their PR's. These individuals become parents for the individuals of the next generation. The probability that an individual will be selected is proportional to its PR.

Once a pair is selected, a check is made to see whether reproduction occurs. Should reproduction occur (according to the probability of crossover, p_c), three offspring are generated according to the following procedure:[13]

$$\begin{aligned} &0.5p_1 + 0.5p_2, \\ &1.5p_1 - 0.5p_2, \\ &-0.5p_1 + 1.5p_2, \end{aligned} \tag{5}$$

where p_1 and p_2 are the two parents. The best two offspring are selected in order to keep the population size constant. The probability of crossover is usually chosen to be between 0.85 and 1. If we use the two individuals of the example of Subsection 3.B, the three offspring created are

n	k	d
2.965	0.030	20.55
2.515	0.078	5.05
3.415	−0.018 (0.087)	36.05 (13.9)

Here the third offspring has values that are not within the specified bounds for k and d. New, random values are then assigned to the parameters that are not within bounds (they are the values within parentheses).

D. Mutation

To ensure variability in the evolutionary process, a chromosome can be subjected to a mutation. A mutation is a random modification of a parameter (gene). A probability of mutation is defined, p_m, usually chosen to be between 0.0001 and 0.1. Should mutation occur, a nonuniform mutation[14] is performed. One of the parameters is modified as follows, after a flip from an unbiased coin,

$$V_i = \begin{cases} V_i + \delta(\mathrm{UB} - V_i) & \text{(heads)} \\ V_i - \delta(V_i - \mathrm{LB}) & \text{(tails)} \end{cases}, \tag{6}$$

where UB is the upper bound of the parameter being mutated and LB is the lower bound. The delta function δ is defined as

$$\delta(y) = y[r(1 - t/T)^B] \tag{7}$$

where r is a random number between 0 and 1, t is the current generation, T is the maximum generation, and B is a parameter that determines the degree of dependence on the actual generation (usually between 1 and 5). From Eq. (7) it can be seen that the amplitude of the mutation decreases as the number of generations increases. This kind of mutation is called nonuniform. Suppose that an individual of values $n = 2.965$, $k = 0.03$, and $d = 20.55$ were to be mutated and that gene 2 had been chosen for mutation. If $t = 10$, $T = 100$, $\mathrm{UB} = 0.1$, $\mathrm{LB} = 0$, $B = 5$, $r = 0.296$, and heads were chosen by the flip of a coin, the value for gene 2 would now be 0.042.

For the problems treated here, elitism[10] had to be used for the GA to converge. Once crossover and mutation have been performed, a check is made to see if individuals of the next generation are better than those of the previous one. The best individual of a generation is called the elite. If the best individual of the next generation, $t + 1$, is better than the best individual of the previous generation, t, then that individual becomes the new elite individual. If the current elite is better than the best individual of the new generation, then the elite is preserved to the next generation. Elitism is done to ensure that good genetic information is not destroyed during reproduction.

Once a new generation is created, the old one is erased. The process of selection, crossover (reproduction), and mutation is repeated until a maximum number of generations is obtained or until the objective function has reached a predetermined value.

4. IMPLEMENTATION

As stated above, the problem now is to inverse the ellipsometric data with an algorithm. Implementing the genetic algorithm is simple. The merit function chosen has some interesting qualities,[8] and the genetic algorithm can easily take advantage of these. Certain parameters must be set, such as the weight exponent, a, the mutation operator, B, the probability of crossover, p_c, the probability of mutation, p_m, the maximum number of generations, T, and the initial population size, η. These parameters are set to standard values and then tweaked to find the right combination that optimizes convergence time. For implementation of the GA to ellipsometric data, the values contained in Table 1 were found to have best performance.

Note that the number of generations is variable. In certain test cases, more generations were necessary to attain total convergence. Usually, GA's converge at 200 generations or less. Certain test cases required 800 generations, owing to the complexity of the problem.

Initial population size is also an issue. Usually a population size of six or seven times the number of variables is enough to diversify the population. For this problem, a population size of 100 was found to be necessary to ensure convergence. Since the value of the objective function varies enormously if one is slightly off the correct solution, a large population is necessary to provide diversification.

Table 1. Optimized Parameters for the GA

Weight (a)	1.2
Mutation operator (B)	3.0
Probability of crossover (p_c)	0.95
Probability of mutation (p_m)	0.05
Number of generations (T)	Variable
Initial population size (η)	100 individuals

Implementing the algorithm is readily done. The only values to be specified are the number of genes (the number of variables) and the boundaries for these genes, and the rest is done by the GA. Computation time is relatively quick on a personal computer, from 5 to 20 s depending on the number of generations and the population size.

5. NUMERICAL ANALYSIS

We performed a number of simulations to test the algorithm. Imaginary samples of different thickness and composition were considered, and random noise was added to the computed ellipsometric data to simulate noise in experimental data. The standard deviation of the added noise was also considered to be the experimental error of the measurements.

The first sample considered is a sample of TiO_2 on glass. The film has the following characteristics: $n = 2.625$, $k = 0.015$, and $d = 417\,\text{nm}$. The substrate has values $n_s = 1.523$ and $k_s = 0$. Light wavelength is 370 nm and the ambient medium is air ($n = 1$). Table 2 shows the computed values of Δ and ψ for four angles of incidence, rounded to the third decimal, and Table 3 shows the same data with random, Gaussian noise added. Noise was added with a random-number generator of specified mean and standard deviation. For the first test case, noise of standard deviations 0.02° and 0.01° was specified, as this is the optimum efficiency of a standard null ellipsometer. Table 3 thus represents the measured values Δ_m and ψ_m. These data were selected for comparison with previously published material.[8]

Table 2. Computed Values of Δ and ψ for a Sample of TiO_2[a] on Glass[b] for Four Incidence Angles[c]

Incidence Angle ϕ	Ellipsometric Angle	
	Δ	ψ
45	170.916	31.981
60	169.064	21.367
70	168.743	8.583
80	357.437	13.043

[a] $n = 2.625$, $k = 0.015$, $d = 417$ nm.
[b] $n_s = 1.523$, $k_s = 0$.
[c] $\lambda = 370$ nm (Ref. 8).

Table 3. Computed Values of Δ and ψ (from Table 2) with Added Random Noise of 0.02° and 0.01° Standard Deviation, Respectively[a]

Incidence Angle ϕ	Ellipsometric Angle	
	Δ_m	ψ_m
45	170.96	31.98
60	169.07	21.37
70	168.73	8.56
80	357.47	13.05

[a] Ref. 8.

The data of Table 3 were incorporated into the GA as the measured values. For each individual (made up of n, k, and d) of the population, values of Δ and ψ were computed and input into the objective function. For this case, the GA converged to values of $n = 2.626$, $k = 0.0153$, and $d = 416.8\,\text{nm}$. The value of the objective function [Eq. (3)] is 0.999. Convergence is obtained after approximately 20 generations. A plot of the convergence for three trial runs is shown in Fig. 1. The boundaries used in this trial were $2.5 < n < 2.9$, $0.001 < k < 0.05$, and $375 < d < 460\,\text{nm}$.

Convergence toward the best solution is not always ensured. Recall that the film thickness is periodic.[15] For this particular problem, the periodicity is approximately 75 nm, which is within the search bounds. Since the initial population is chosen at random, it is possible that the best individual is one that is closest to the other periodic solution. Given that the best individual is also reproduced more often, the algorithm may converge to the other solution. Increasing the population size and reducing the weight of the best individuals solves this problem.

Different levels of noise were considered to test the algorithm. This is reflected in the data of Tables 4 and 5. Table 4 shows data for the angles Δ and ψ where random noise of standard deviations 0.5° and 0.25°, respectively, were added.

Results obtained with the data from Table 4 show that precision decreases as error increases. The GA converges to values of $n = 2.665$, $k = 0.0206$, and $d = 409.8\,\text{nm}$. The objective function has a value of 1.14. Now we take another interesting case, where the error is known and is different for each measurement. Table 5 contains these data.

In this case, the GA converges to values of $n = 2.617$, $k = 0.0134$, and $d = 418.5\,\text{nm}$, with an objective function value of 1.15. Should the error on the measurements be unknown, a very possible real situation, one can simply assume that $\epsilon^{\Delta} = 2$ and $\epsilon^{\psi} = 1$, since the error on Δ is usually twice that of ψ (Ref. 8). For this interesting case, with the data of Table 3, the GA converges to values of $n = 2.615$, $k = 0.0137$, and $d = 418.8\,\text{nm}$. The objective function has a value of 0.019, but this value has no real significance, since the errors chosen here have no physical basis. We can say that the error has been overestimated, since the objective function is $\ll 1$, and we know that this is true since the errors were 0.02° and 0.01°.

To test the algorithm further, consider another film–substrate system.[8] In this case, the film has properties of $n = 3.410$, $k = 2.630$, and $d = 10\,\text{nm}$. The substrate is $n_s = 3.858$ and $k_s = 0.018$. Measurements are at $\lambda = 632.8\,\text{nm}$. Values of Δ and ψ were generated and rounded to the third decimal. For the first test, we considered the data to be exact, meaning $\epsilon^{\Delta} = 1$ and $\epsilon^{\psi} = 1$ (no weighting is done by the errors). The GA converges to $n = 3.412$, $k = 2.623$, and $d = 10.03\,\text{nm}$. The problem is much more complex, as the GA converges only after 800 generations for all trials. The objective function is 0.999. In this case, the boundaries for the three unknowns are $3.0 < n < 4.0$, $1.5 < k < 3.5$, and $1.0 < d < 30\,\text{nm}$. To simulate a more real situation, the data were rounded to the second decimal point and the errors were considered to be $\epsilon^{\Delta} = 0.02°$ and $\epsilon^{\psi} = 0.01°$.

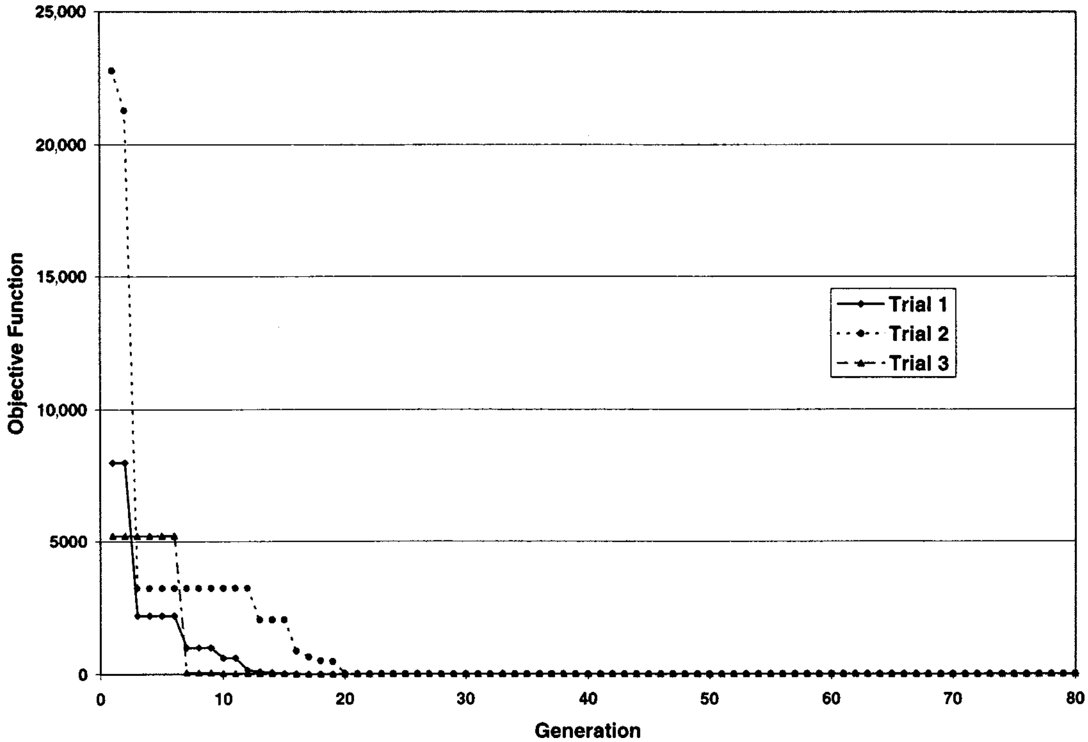

Fig. 1. Evolution of the objective function as the generation increases for three different trials (Δ_m, ψ_m from Table 3). Note that convergence is achieved after approximately 20 generations.

Table 4. Computed Values of Δ and ψ (from Table 2) with Added Random Noise of 0.5° and 0.25° Standard Deviation, Respectively[a]

Incidence Angle ϕ	Ellipsometric Angle Δ	Ellipsometric Angle ψ
45	171.21	32.40
60	169.22	21.47
70	169.19	8.34
80	358.51	12.97

[a] Ref. 8.

Table 5. Data Obtained by Adding to the Data of Table 2 Gaussian Noise with Different Standard Deviations[a]

Incidence Angle	Ellipsometric Angle		Standard Deviation	
ϕ	Δ_m	ψ_m	σ^{Δ}	σ^{ψ}
45	170.95	31.93	0.22	0.06
60	168.73	21.48	0.38	0.06
70	169.23	8.42	0.52	0.20
80	357.14	13.01	0.44	0.12

[a] Ref. 8. The measurement errors ϵ^{Δ} and ϵ^{ψ} were chosen to be the same as the standard deviations.

The algorithm still converges, at a point of $n = 3.369$, $k = 2.764$, and $d = 9.40$ nm. The objective function has a value of 1.14.

Another test was done on a third film–substrate system, where the film is $n = 1.6$, $k = 0.5$, and $d = 25$ nm. The substrate is $n_s = 3.85$ and $k_s = 0.02$, and measurements are made at $\lambda = 632.8$ nm. This time, data are generated for only two angles of incidence, 50° and 70°, and no noise is added to the data. The GA converges to $n = 1.599$, $k = 0.505$, and $d = 24.94$ nm, with an objective function value of 6.4×10^{-5} after 200 generations.

A final simulation was done with a film of $n = 2.2$, $k = 0.22$, and $d = 10$ nm on a substrate of $n_s = 4.05$, and $k_s = 0.028$ with $\lambda = 546.1$ nm. Data were computed and then noise was added with standard deviation of 0.02° and 0.01°. Two trials were done, one with two angles of incidence and one with four angles of incidence. Table 6 shows the data with and without added noise.

The GA converges after 500 generations to $n = 1.98$, $k = 0.34$, and $d = 9.59$ nm. The objective function is 1.48. Table 7 shows data from the same substrate as in Table 6, this time with measurements at four angles instead of two. Again, noise was added to the data to simulate experimental errors. Noise of standard deviation 0.02° and 0.01° is added.

One would think that the algorithm would be able to find a better solution with four than with only two measurements. The GA converges to $n = 1.8$, $k = 0.75$, and $d = 8.32$ nm, where the objective function is 1.99. This shows that taking more measurements does not necessarily improve precision, because of measurement errors.

6. RESULTS AND DISCUSSION

Results obtained by the GA are similar to those obtained by the modified downhill simplex algorithm[8] and the re-

Table 6. Data Generated for a Film of $n = 2.2$, $k = 0.22$, and $d = 10$ nm on a Substrate of $n_s = 4.05$ and $k_s = 0.028$ with $\lambda = 546.1$ nm

Incidence Angle	Generated Data		Noise Added	
ϕ	Δ	ψ	Δ_m	ψ_m
50	172.393	31.959	172.39[a]	31.96[b]
70	145.506	12.886	145.49[a]	12.89[b]

[a] Added noise of 0.02° standard deviation.
[b] Added noise of 0.01° standard deviation.

Table 7. Data Generated for a Film of $n = 2.2$, $k = 0.22$, and $d = 10$ nm on a Substrate of $n_s = 4.05$ and $k_s = 0.028$ with $\lambda = 546.1$ nm

Incidence Angle	Generated Data		Noise Added	
ϕ	Δ	ψ	Δ_m	ψ_m
45	174.267	34.835	174.30[a]	34.84[b]
50	172.393	31.959	172.41[a]	31.96[b]
70	145.506	12.886	145.48[a]	12.88[b]
80	32.836	13.417	32.84[a]	13.40[b]

[a] Added noise of 0.02° standard deviation.
[b] Added noise of 0.01° standard deviation.

sults of other inversion techniques. The GA was tested on a number of test cases, and results obtained are quite good. Also, the genetic algorithm does not require a starting value, so evaluation of an unknown sample is possible. The only problem to be aware of is the periodicity of the sample thickness. If the layer periodicity is within the search bounds, the algorithm may converge to the periodic thickness, and although the answer is mathematically correct, it does not represent the actual physical thickness. Computation times of the GA vary, but a trial of 200 generations with a population of 100 requires approximately 16 s to compute on a P-II 400-MHz personal computer. Some problems with GA's can be solved without the use of elitism, but in this case elitism was necessary to ensure convergence. Without elitism, the algorithm fluctuated wildly and did not converge to a minimum.

7. CONCLUSION

A new method has been proposed to invert ellipsometric data. A genetic algorithm has been applied to data to determine the optical properties of thin films. The GA is easy to implement and does not require much computation time. A number of test cases have been done to simulate real situations. The GA performs well, converging to the best possible values. The GA can be used in a number of different situations, such as cases with different numbers of measurements or different errors on the measurements.

The GA can easily be applied to systems with more than one layer, since the method does not require evaluation of derivatives and can easily be adapted to any number of variables.

ACKNOWLEDGMENTS

The authors thank Noyan Turkkan for his help with genetic algorithms, the reviewers for their helpful comments, and the Natural Sciences and Engineering Research Council of Canada for financial support.

All correspondence should be addressed to Roger Boudreau, École de génie, Université de Moncton, Moncton, NB, Canada E1A 3E9.

REFERENCES

1. R. M. A. Azzam and N. M. Bashara, *Ellipsometry and Polarized Light* (North-Holland, New York, 1996).
2. F. L. McCrackin and J. P. Colson, "Computational techniques for the use of the exact Drude equations in reflection problems," in "Ellipsometry in the measurement of surfaces and films," Natl Bur. Stand. Misc. Publ. **256**, 61–82 (1964).
3. A. R. Reinberg, "Ellipsometer data analysis with a small programmable desk calculator," Appl. Opt. **11**, 1273–1274 (1972).
4. T. Easwarakhanthan, C. Michel, and S. Ravelet, "Numerical method for the ellipsometric determination of optical constants and thickness of thin films with microcomputers," Surf. Sci. **197**, 339–345 (1988).
5. F. K. Urban, "Ellipsometry algorithm for absorbing films," Appl. Opt. **32**, 2339–2344 (1993).
6. J.-P. Drolet, S. C. Russev, M. I. Boyanov, and R. Leblanc, "Polynomial inversion of the single transparent layer problem in ellipsometry," J. Opt. Soc. Am. A **11**, 3284–3291 (1994).
7. G. E. Jellison, "Use of the biased estimator in the interpretation of spectroscopic ellipsometry data," Appl. Opt. **30**, 3354–3360 (1991).
8. S. Bosch, F. Monzonís, and E. Masetti, "Ellipsometric methods for absorbing layers: a modified downhill simplex algorithm," Thin Solid Films **289**, 54–58 (1996).
9. J. H. Holland, *Adaptation in Natural and Artificial Systems* (University of Michigan Press, Ann Arbor, Mich., 1975).
10. D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning* (Addison-Wesley, Reading, Mass., 1989).
11. T. Eisenhammer, M. Lazarov, M. Leutbecher, U. Schöffel, and R. Sizmann, "Optimization of interference filters with genetic algorithms applied to silver-based heat mirrors," Appl. Opt. **32**, 6310–6315 (1993).
12. L. Davis, *Handbook of Genetic Algorithms* (Van Nostrand Reinhold, New York, 1991).
13. A. H. Wright, "Genetic algorithms for real parameter optimization," in *Foundation of Genetic Algorithms*, G. J. E. Rawlins, ed. (Morgan Kaufmann, San Mateo, Calif., 1991), pp. 205–218.
14. Z. Michalewicz, *Genetic Algorithms* (Springer-Verlag, New York, 1992).
15. S. Bosch, J. Perez, and A. Canillas, "Numerical algorithm for spectroscopic ellipsometry of thick transparent films," Appl. Opt. **37**, 1177–1179 (1998).