

Bayesian inference analysis of ellipsometry data

N. P. Barradas*

School of Electronic Engineering, Information Technology and Mathematics, University of Surrey, Guildford, Surrey GU2 5XH, United Kingdom

J. L. Keddie[†] and R. Sackin

School of Physical Sciences, University of Surrey, Guildford, Surrey GU2 5XH, United Kingdom

(Received 8 September 1998)

Variable angle spectroscopic ellipsometry is a nondestructive technique for accurately determining the thicknesses and refractive indices of thin films. Experimentally, the ellipsometry parameters ψ and Δ are measured, and the sample structure is then determined by one of a variety of approaches, depending on the number of unknown variables. The ellipsometry parameters have been inverted analytically for only a small number of sample types. More general cases require either a model-based numerical technique or a series of approximations combined with a sound knowledge of the test sample structure. In this paper, the combinatorial optimization technique of simulated annealing is used to perform least-squares fits of ellipsometry data (both simulated and experimental) from both a single layer and a bilayer on a semi-infinite substrate using what is effectively a model-free system, in which the thickness and refractive indices of each layer are unknown. The ambiguity inherent in the best-fit solutions is then assessed using Bayesian inference. This is the only way to consistently treat experimental uncertainties along with prior knowledge. The Markov chain Monte Carlo algorithm is used. Mean values of unknown parameters and standard deviations are determined for each and every solution. Rutherford backscattering spectrometry is used to assess the accuracy of the solutions determined by these techniques. With our computer analysis of ellipsometry data, we find all possible models that adequately describe that data. We show that a bilayer consisting of a thin film of poly(styrene) on a thin film of silicon dioxide on a silicon substrate results in data that are ambiguous; there is more than one acceptable description of the sample that will result in the same experimental data. [S1063-651X(99)02905-0]

PACS number(s): 02.70.Lq, 78.20.Ci, 78.20.Bh, 02.60.Ed

I. INTRODUCTION

Ellipsometry is a fast accurate technique for measuring the optical constants, interfacial roughness, and thicknesses of thin films. The technique has become widespread over the last 30 years in a diverse range of fields. Two recent reviews [1,2] highlight examples of ellipsometry applications. These include the determination of glass transition temperatures in polymer thin films [3], thin film swelling [4], adsorption of small molecules at solid/liquid interfaces [5,6], the characterization of Langmuir-Blodgett films [7] and the determination of damage depth profiles from ion implantation in silicon wafers [8]. For a given sample, ellipsometry measures ellipticity ρ , which is written as

$$\rho = \frac{R_p}{R_s} = \tan \psi e^{i\Delta}, \quad (1)$$

where R_p and R_s are the Fresnel reflection coefficients, with p denoting the plane of reflection and s denoting the plane perpendicular to it [9]. The Fresnel coefficients are dependent on experimental parameters: the angle-of-incidence of

light, φ_0 (conventionally measured from the sample surface normal), and the wavelength of radiation, λ . The Fresnel coefficients are also functions of material parameters: complex refractive indices of each of the components, N , and each of the layer thicknesses, d . The parameter Δ is the change in phase difference between the p and s components caused by reflection, while ψ is the ratio of the amplitude ratios of the p and s light components before and after reflection. The ellipticity is measured by the analysis of the elliptically polarized light reflected from a flat, smooth sample surface. The ratio can be described algebraically by an expression derived from the Fresnel coefficients for n layers on a semi-infinite substrate, where n is any integer number [9]. Starting with (ψ, Δ) pairs obtained at known φ_0 and λ , one can invert this expression, under certain circumstances outlined below, to find values for the unknown parameters, such as N , d , and the roughness of each layer in a sample. Only a few specific cases have as yet been inverted analytically. Drolet *et al.* [10] summarized these structures as (1) a single layer with unknown complex refractive index on a known substrate; (2) a substrate with two layers and with one layer thickness being unknown; (3) multilayer systems with any one unknown layer thickness; (4) multilayer systems with unknown substrate complex refractive index; (5) a symmetric system of one layer of unknown thickness and real refractive index embedded in two identical phases having a real index; and (6) an optically absorbing layer on a substrate with a complex refractive index, and with the thickness of the layer being unknown.

*Permanent address: Instituto Tecnológico Nuclear, E.N. 10, 2685 Sacavém, Portugal.

[†]Author to whom correspondence should be addressed. Electronic address: j.keddie@surrey.ac.uk

Cases (1)–(5) require measurements on *one* substrate at *one* wavelength and *one* angle. Case (6) is somewhat different in that it requires *two* different angles of incidence on *two* different substrates. It should be noted that inherent correlation exists between measurements at two different values of φ_0 and λ . Hence, increasing the number of measurements does not necessarily increase the amount of information about a sample. Moreover, the use of multiple λ values potentially introduces more unknown material parameters due to optical dispersion and absorption. In some cases, however, the number of unknowns can be reduced by using analytical or empirical expressions to describe a material’s optical properties [11,12].

A number of numerical inversion methods have been developed which are suitable for different problems. There are at least three categories of methods, as outlined below.

(1) There are exact numerical methods suitable for when the number of unknowns is equal to the number of ψ and Δ measurements. The Reinberg method [13] is most often cited for single-angle, single-wavelength ellipsometry to determine the unknown refractive index and thickness of a thin nonabsorbing film on a reflecting substrate with known optical constants. A similar method is the functional-link neural network approach of Park *et al.* [14]. Other techniques are cited by Drolet *et al.* [10].

(2) When the number of unknown variables is greater than the number of ψ and Δ data points, an exact solution can obviously not be written. Multiparameter fitting methods, of which the Levenberg-Marquardt [15] algorithm is commonly used, are suitable for data from variable angle spectroscopic ellipsometry (VASE) (a technique where many φ_0 and λ are used). Such an algorithm is used to minimize the difference between the experimental (ψ, Δ) spectra and the simulated spectra generated from the Fresnel equations [12,9]. Models can be built to include any number of layers, with complex refractive indices, on any substrate and in any ambient medium. Other physical properties such as biaxiality [16] and surface roughness can also be included in the model. A related technique is the backpropagation neural network method of Fried and Masa [17], which is trained to recognize characteristics of (ψ, Δ) spectra.

(3) There exist inversion-after-approximation schemes, such as that described by Charmet and de Gennes [18] for multiple-angle single-wavelength ellipsometry. This technique can determine an arbitrary refractive index profile over depths much greater than $\lambda/4\pi$.

The preceding methods, although for the most part quick to perform, are all limited when little or nothing is known about the test sample. The exact numerical methods are good only for specific, very simple samples. The multiparameter fitting methods only globally minimize the difference between the simulated and experimental spectra if the initial guess solution is close to the global minima. Otherwise, local minima solutions are given. The neural network schemes presented in the literature have, so far, only been trained successfully for a limited set of solutions. The Charmet–de Gennes method [18] can only give results for specific sample types over certain depth scales. Knowledge of the substrate and ambient refractive indices is also required. There is clearly a need for a *model-independent* technique when a sample is presented in which the structure is not well known.

The use of VASE has become widespread for the study of complex samples, as it is the most powerful ellipsometry technique. In any approach to the analysis of VASE data, the aim is to determine a unique physical description, or model, of the unknown sample. Except in the case of exact data inversion, however, there is a lingering question of the uniqueness of the model obtained. Hence, prior to fitting the data to a model, one should assess the solution space for ambiguities. That is, one should find the number of exact solutions that exist for a given set of (ψ, Δ) pairs. A study of ambiguity in solution space has recently been presented by Polovinkin and Svitashva [19] using a “step-by-step movement” numerical method to search solution space for data from a single unknown layer deposited on a semi-infinite substrate. However, expanding this technique for an n -unknown problem would not be trivial. Unless the increments used in the steps are infinitesimally small, there is a finite possibility that such a search will not find all solutions to the problem. The search might miss a description of the sample that is the “true” description. We have therefore developed an alternative method for assessing ambiguities.

In this paper, we use the simulated annealing (SA) algorithm to perform multiparameter least-square fits to ellipsometry data. Simulated annealing is a global optimization algorithm designed to find the absolute minimum (or maximum) of any given function [20–22]. It is completely general in that it entails in principle no restrictions on the function to be minimized. In the case of ellipsometry data analysis, no assumptions need to be made about the sample’s physical properties. SA has solved previously intractable problems such as the traveling salesman problem [23], and it is widely applied in fields ranging from ion beam analysis [24–29] to natural language processing [30]. We show here that SA finds solutions that correctly reproduce VASE data. However, since it is a stochastic technique, if more than one sample structure can fit the data, SA will randomly find only *one* of the possible structures. To overcome this limitation, we have also applied the Bayesian inference, which is the *only* way to consistently treat incomplete and noisy data when additional prior information is known. It is realised using the Markov chain Monte Carlo (MCMC) algorithm [31–33]. The MCMC algorithm explores the whole parameter space, and is therefore able to find each and every solution that is consistent with the data. Beneficially, it provides confidence limits on the solutions obtained. It has already been successfully applied to other techniques [34–37]. We apply SA and MCMC techniques to both theoretically generated and experimental ellipsometry data.

II. THEORY OF ELLIPSOMETRY: DERIVATION OF THE FRESNEL REFLECTION COEFFICIENTS

The central equation of ellipsometry is given by Eq. (1). In this section, we present expressions for the Fresnel coefficients [12]. For an m -layer system on a substrate (in which m is a positive integer and the substrate is the $m + 1$ th layer), Snell’s law states

$$N_0 \sin \varphi_0 = N_1 \sin \varphi_1 = \dots = N_j \sin \varphi_j = \dots = N_{m+1} \sin \varphi_{m+1}, \quad (2)$$

where N_j is the refractive index of the j th layer and φ_j represents the angle between the direction of propagation in the j th layer and the perpendicular to the plane of the layer's interfaces. N_0 is the ambient refractive index. For a system where the thickness of the j th layer is d_j , the Fresnel reflection coefficients are given by

$$R_p = \frac{S_{21p}}{S_{11p}} \quad (3)$$

and

$$R_s = \frac{S_{21s}}{S_{11s}}, \quad (4)$$

where

$$\mathbf{S}_{p,s} = \begin{bmatrix} S_{11p,s} & S_{12p,s} \\ S_{21p,s} & S_{22p,s} \end{bmatrix}. \quad (5)$$

The matrix \mathbf{S} is given by

$$\mathbf{S}_{p,s} = \mathbf{I}_{01p,s} \mathbf{L}_1 \mathbf{I}_{12p,s} \mathbf{L}_2 \cdots \mathbf{I}_{(j-1)p,s} \mathbf{L}_j \cdots \mathbf{L}_m \mathbf{I}_{m(m-1)p,s}, \quad (6)$$

with

$$\mathbf{I}_{abp,s} = \begin{bmatrix} \frac{1}{t_{ab}} & \frac{r_{ab}}{t_{ab}} \\ \frac{r_{ab}}{t_{ab}} & \frac{1}{t_{ab}} \end{bmatrix}, \quad (7)$$

$$\mathbf{L}_a = \begin{bmatrix} e^{i\beta_a} & 0 \\ 0 & e^{-i\beta_a} \end{bmatrix}, \quad (8)$$

and

$$\beta_a = \left(\frac{2\pi d_a N_a}{\lambda} \right) \cos \varphi_a. \quad (9)$$

r_{ab} is the amplitude reflection coefficient at the interface between substance a and b , and t_{ab} is the amplitude transmission coefficient of the ab interface. By convention, the ambient substance is designated as medium 0, and the substrate medium is designated with the highest number. The reflection and transmission coefficients are expressed by

$$r_{ab} = \frac{\nu_a - \nu_b}{\nu_a + \nu_b} \quad (10)$$

and

$$t_{ab} = \frac{2\nu_a}{\nu_a + \nu_b}, \quad (11)$$

where, for the s -polarized component,

$$\nu_a = N_a \cos \varphi_a, \quad (12)$$

and, for the p -polarized component,

$$\nu_a = \frac{N_a}{\cos \varphi_a}. \quad (13)$$

Other expressions can be included to incorporate features such as surface roughness and anisotropic refractive indices [16]. In this paper, we present analysis of a bilayer (i.e., two adjacent parallel films) at the interface of a semi-infinite substrate and an ambient medium. In this case, the Fresnel coefficients can be written as [9,12]

$$R_{p,s} = \frac{(r_{01p,s} + r_{12p,s} e^{-i2\beta_1}) + (r_{01p,s} r_{12p,s} + e^{-i2\beta_1}) r_{23p,s} e^{-i2\beta_2}}{(1 + r_{01p,s} r_{12p,s} e^{-i2\beta_1}) + (r_{12p,s} + r_{01p,s} + e^{-i2\beta_1}) r_{23p,s} e^{-i2\beta_2}}. \quad (14)$$

When analyzing dielectric materials, it is convenient to use an analytical expression to describe the optical dispersion, and thereby reduce the number of unknown parameters. According to the Cauchy dispersion model [12],

$$N_j(\lambda) = na_j + \frac{nb_j}{\lambda^2} + \frac{nc_j}{\lambda^4} + \dots, \quad (15)$$

where na_j , nb_j , and nc_j are constants for the j th layer, and λ , by convention, is given here in units of μm . The nc_j terms and above are small and can be neglected. The aim of an ellipsometry inversion algorithm is to find values for the unknown N_j and d_j that produce the measured (ψ, Δ) spectra. In some cases, there might be more than one combination of N_j and d_j that are solutions within the experimental errors of the data.

III. SIMULATED ANNEALING ALGORITHM

Simulated annealing is based on an analogy with the mathematics of the thermal annealing of crystals in which defects from a crystal are removed by melting it and subsequently cooling it down very slowly. In the annealing process, from the current state i of the system with energy E_i , another state j with energy E_j is generated by a random process, in which the state i is slightly altered. If the energy decreases in the transition, that is, if $\Delta E = E_j - E_i < 0$, the system is taken to be in the new state j . If the energy increases, then the transition has a certain probability given by the Boltzmann factor

$$P(i \rightarrow j) = \exp(-\Delta E/k_B T) \quad \text{if } \Delta E > 0, \quad (16)$$

TABLE I. Original structure used to simulate ellipsometry data and the corresponding main features of the MCMC solutions 1, 2, and 3 for this structure upon analysis of the simulated data. The semi-infinite substrate was chosen to have $N=1.5$, and the ambient medium had $N=1.0$ (corresponding to a vacuum).

Solution	d_1 (nm)	na_1	nb_1	d_2 (nm)	na_2	nb_2
original structure	10.0	1.25	0.0015	55.0	1.45	0.0030
1	≈ 10	broad distribution	broad distribution	≈ 50	≈ 1.45	≈ 0.003
2	≈ 15	broad distribution	broad distribution	≈ 80 , tail down to 20	≈ 1.48	broad distribution
3	≈ 50	≈ 1.4	≈ 0.005	≈ 0	undetermined	undetermined

where T is the absolute temperature of the system, and k_B is the Boltzmann constant. If the initial temperature of the system is high enough, and if at each temperature enough time is allowed for the system to reach thermodynamic equilibrium, and if the cooling rate is slow enough, then at $T=0$ K the crystal is guaranteed to be in a state of minimum energy. Note that if the system is degenerate (i.e., more than one state corresponds to the minimum energy) at $T=0$ K the system will be in any one of the minimum-energy degenerate states.

In the SA analogy, E is represented by any objective function $f \equiv f(\mathbf{x})$ to be minimized, and the state of the system is defined as \mathbf{x} . In ellipsometry, \mathbf{x} is the thickness, complex refractive index and other physical properties of all layers in the sample. A state transition is defined as the generation of a new structure given the previously calculated state by randomly changing d_j and N_j by some amount. The probability P of accepting a transition from state \mathbf{x} to state \mathbf{y} is given by the so-called Metropolis criterion [38]

$$P(\mathbf{x} \rightarrow \mathbf{y}) = \min\{\exp(-\Delta\chi^2/T), 1\}, \quad (17)$$

where $f \equiv \Delta\chi^2$ is the change in χ^2 due to the transition, and T is a control parameter. In the computer implementation of the algorithm, $P(\mathbf{x} \rightarrow \mathbf{y})$ is calculated and compared with a random number $r \in [0,1]$, and the transition is accepted if $P(\mathbf{x} \rightarrow \mathbf{y}) > r$. At high values of the control parameter T , practically all the transitions are accepted, corresponding in the analogy to a liquid state with high entropy. As T decreases, the probability of transitions decreases. At very small values of T , only transitions that lead to a decrease in χ^2 are accepted. In SA, T is initialized at some high value T_0 that allows almost all transitions to be accepted. Then T is decreased slowly, according, for instance, to

$$T_{i+1} = kT_i \quad \text{where } 0 < k < 1. \quad (18)$$

At each value of T, L_M transitions are proposed. The succession of all the accepted states is called a Markov chain. During the SA process several Markov chains are computed, one for each value of the control parameter. The values of T_0, k , and L_M define what is called the cooling schedule. For sufficiently high values of these three parameters, it can be mathematically proven [20–22] that the global minimum of the objective function is found. Given a sufficient cooling

schedule, SA is thus *guaranteed* to find the best-fit model to ellipsometry data. Practically, such a procedure of SA would lead to extremely long calculation times. A reasonable cooling schedule, which leads to a high quality solution as opposed to the best one, must therefore be used in most situations.

SA, while very successful in a wide range of problems, has a major shortcoming: it returns one state of the system corresponding to the global minimum χ^2 found, without any indication of fit error. Even so, it is superior to other algorithms that are prone to give only a local minimum and depend on starting values. It is highly desirable to obtain *all*

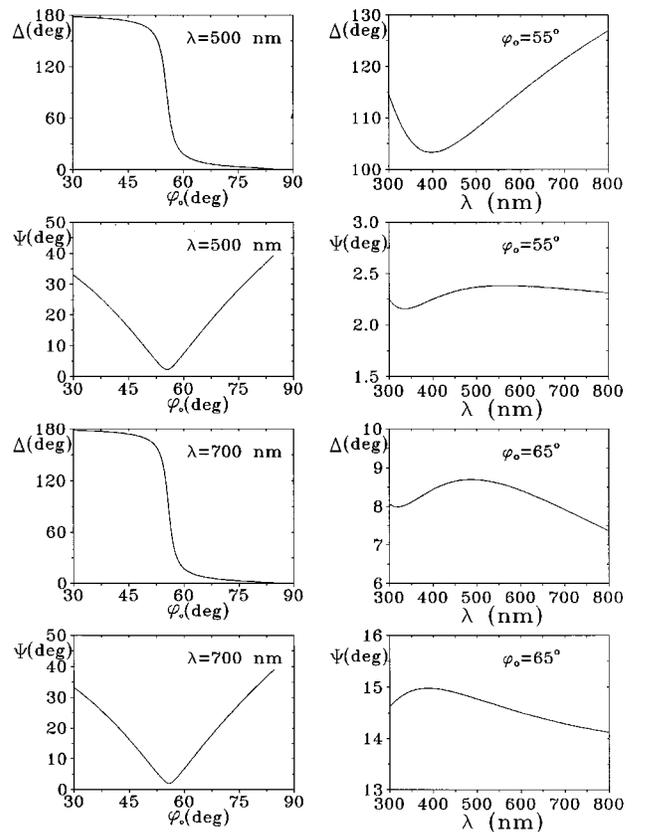


FIG. 1. Simulated spectroscopic ellipsometry scans generated for a bilayer at a vacuum-substrate interface. Values of d_j and N_j are given in Table I. Simulated plots of Ψ and Δ are shown as functions of ϕ_0 for (a) $\lambda = 500$ nm and (b) $\lambda = 700$ nm, and as functions of λ for (c) $\phi_0 = 55^\circ$ and (d) $\phi_0 = 65^\circ$.

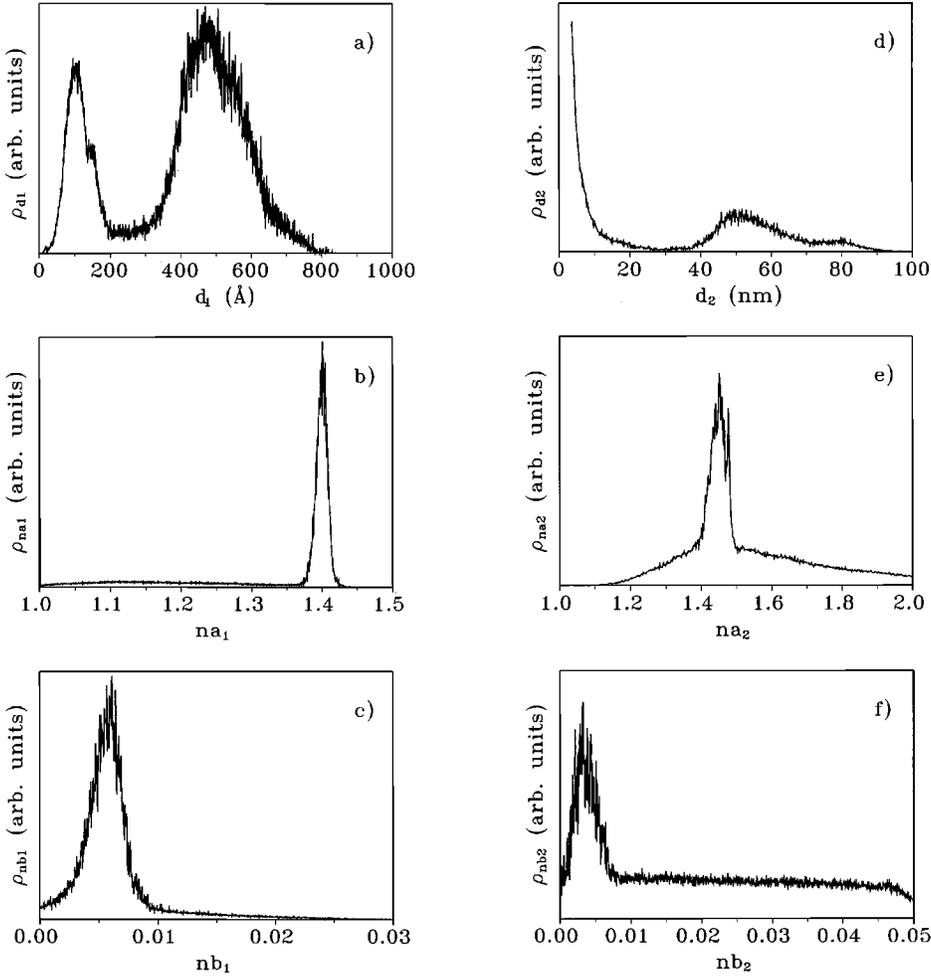


FIG. 2. Density of states obtained with a MCMC analysis of simulated data that was generated from the structure described in Table I, for each of six “unknown” parameters: (a) d_1 , (b) na_1 , (c) nb_1 , (d) d_2 , (e) na_2 , and (f) nb_2 .

possible solutions with corresponding confidence intervals. A means to this end are described in Sec. IV.

IV. BAYESIAN INFERENCE AND THE MARKOV CHAIN MONTE CARLO ALGORITHM

Suppose one wishes to study \mathbf{x} , about which some *a priori* information I exists which can be expressed in terms of the conditional probability $p(\mathbf{x}|I)$, that is, the probability of \mathbf{x} given I . This is known as the *prior distribution*. It is the knowledge one has about \mathbf{x} before the experiment under consideration is done, and can include information coming from other sources, for instance complementary experiments. Suppose also that some experimental observations $d \equiv d(\mathbf{x})$ exist, which depend on the parameters x in a known way. The knowledge of the dependence of the observations upon \mathbf{x} is then the conditional probability $p(d|\mathbf{x}I)$, the so-called *likelihood function*. It describes how probable it is to obtain a certain experimental result given well-known parameters. In ellipsometry, that corresponds to calculating theoretical angular or wavelength scans from known layer thicknesses and refractive indices. Bayes’ theorem describes how much the experimental observations alter the original beliefs about the parameter \mathbf{x} :

$$p(\mathbf{x}|dI) = p(d|\mathbf{x}I)p(\mathbf{x}|I)/p(d|I). \quad (19)$$

The probability density function $p(\mathbf{x}|dI)$ is called the *posterior distribution*. It is knowledge about \mathbf{x} gained from experiments and any other measurements. It is possible to use $p(\mathbf{x}|dI)$ to calculate the mean solution $\langle \mathbf{x} \rangle$, as well as confidence intervals given by the standard deviation $\sigma(\mathbf{x})$ of the solution. If the problem is multimodal, that is, if there is more than one solution, means and errors for each solution can be calculated, thereby obtaining all solutions of the problem.

As $p(d|I)$ is independent of \mathbf{x} , we can treat it as a normalization constant. In the general case, any thickness and refractive index values are possible, so $p(\mathbf{x}|I)$ could also be constant. However, in ellipsometry problems, it is more often the case that previous information about the system is available from other sources, such as the refractometry of the substrate, and this term represents these constraints. Hence in these cases it is not convenient to use maximum-entropy’s uninformative prior distribution [34]. Finally, $p(d|\mathbf{x}I)$ is taken to depend on χ^2 through

$$p(d|\mathbf{x}I) \propto \exp(-\chi^2/2), \quad (20)$$

where the χ^2 function is defined in the usual way [39]:

$$\chi^2 \equiv \sum_q [(Y_{\text{expt}}^q - Y_{\text{theor}}^q(\mathbf{x}))/\sigma_q]^2, \quad (21)$$

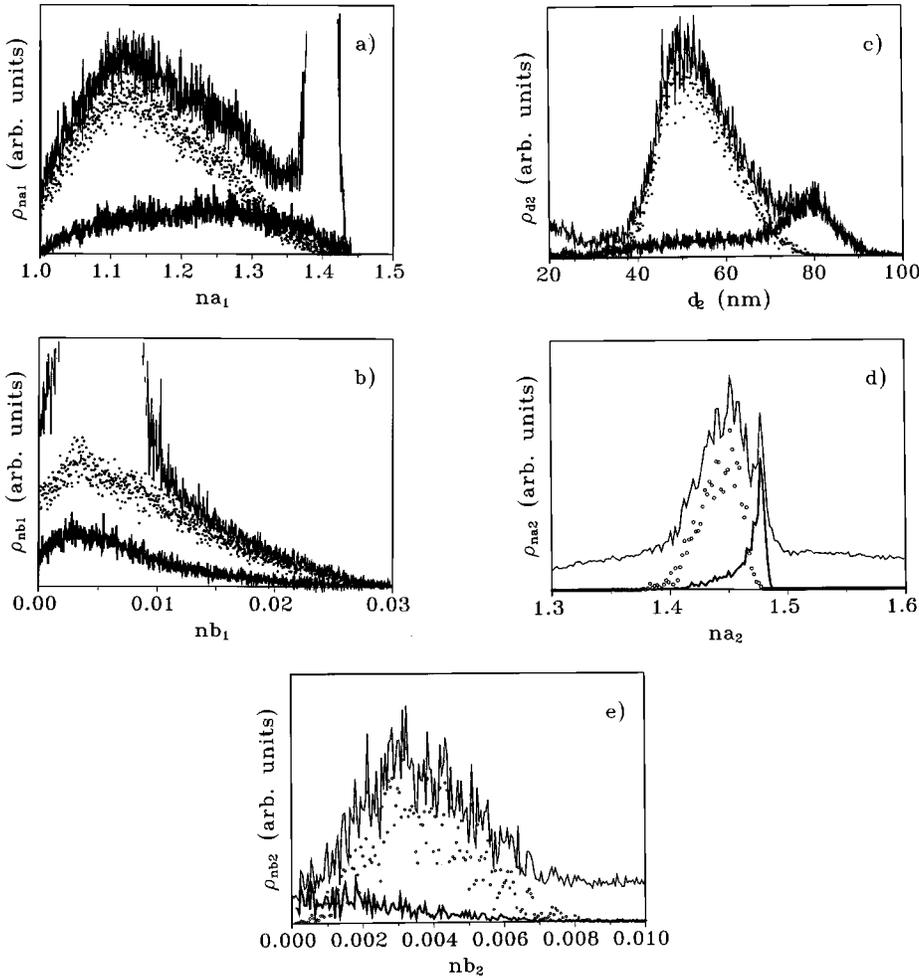


FIG. 3. Density of states obtained with a MCMC analysis for (a) na_1 , (b) nb_1 , (c) d_2 , (d) na_2 , and (e) nb_2 for simulated data. The dots are the partial densities for solution 1, the lower solid lines are the partial densities for solution 2, and the upper solid lines are for solution 3.

where Y_{expt}^q and Y_{theor}^q are the experimental and theoretical values, respectively, and σ_q is the experimental error at point q . Other sources of error, such as uncertainty in the values of φ_0 or λ , can also be included.

A sample can be drawn from equation (19) using the MCMC sampling method, in particular, the Metropolis-Hastings algorithm [38] based on the Metropolis criterion. Consider a Markov chain $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n, \dots, \mathbf{x}_m$, with

$$p(\mathbf{x}_i | \mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{i-1}) = p(\mathbf{x}_i | \mathbf{x}_{i-1}), \quad (22)$$

that is, the probability that the i th member of the chain is \mathbf{x}_i depends only on the previous element of the chain. That probability is determined by a random distribution $q(\mathbf{x}_i, \mathbf{y})$. The Markov chain is then generated by proceeding from \mathbf{x}_i to \mathbf{x}_{i+1} by considering a candidate \mathbf{y} generated with the random distribution $q(\mathbf{x}_i, \mathbf{y})$. The candidate \mathbf{y} is then accepted (that is, it becomes \mathbf{x}_{i+1}) with probability $P(\mathbf{x} \rightarrow \mathbf{y})$ according to the generalized Metropolis criterion

$$P(\mathbf{x} \rightarrow \mathbf{y}) = \min\left\{\frac{\pi(\mathbf{y})q(\mathbf{y}, \mathbf{x})}{\pi(\mathbf{x})q(\mathbf{x}, \mathbf{y})}, 1\right\}, \quad (23)$$

where the acceptance function $\pi(\mathbf{x})$ is a function of \mathbf{x} (see below). After running the Markov chain until equilibrium has been reached [until the probability of the system being in the state \mathbf{x} is given by $\pi(\mathbf{x})$], further $(m-n)$ iterations are calculated; the Markov chain $\mathbf{x}_n, \dots, \mathbf{x}_m$ so generated constitutes a sample from the acceptance distribution $\pi(\mathbf{x})$. This

sample is an empirical distribution which, if large enough, reflects all the properties of π without needing to evaluate it directly. The acceptance function is defined as

$$\pi(\mathbf{y}) = p(d|\mathbf{x}I)p(\mathbf{x}|I). \quad (24)$$

If the transition distribution q is chosen such that it is symmetric, i.e.,

$$q(\mathbf{x}, \mathbf{y}) = q(\mathbf{y}, \mathbf{x}), \quad (25)$$

then generating candidate \mathbf{x} from the current element \mathbf{y} is equally probable, and vice versa. The acceptance criterion to generate the Markov chain becomes

$$P(\mathbf{x} \rightarrow \mathbf{y}) = \min\{\exp(-\Delta\chi^2/2), 1\}. \quad (26)$$

The Markov chain so generated is then a sample of $p(\mathbf{x}|dI)$, which means that it reflects *all* the information over \mathbf{x} that can be obtained from the experimental data, taking into account the experimental uncertainties as well as any previous system information. Not only can the averages and standard deviations of the thicknesses and refractive indices be calculated, but ambiguous problems can also be conveniently treated. If there is more than one solution that fits the data correctly, $p(\mathbf{x}|dI)$ will be multimodal.

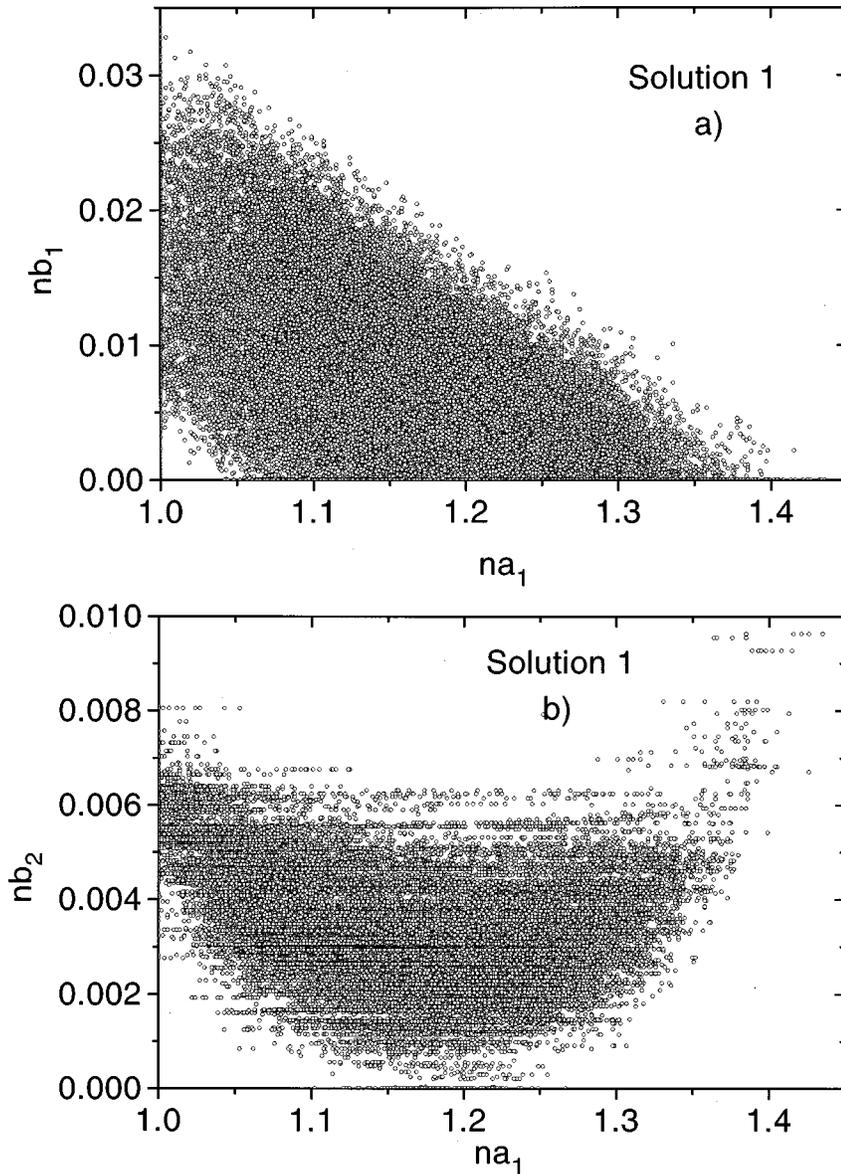


FIG. 4. Solution 1 from a MCMC analysis of simulated ellipsometry data (based on the structure described in Table I) shown as (a) nb_1 vs na_1 and as (b) nb_2 vs na_1 . Each dot in the figures represents one element x_i of the Markov chain.

V. EXPERIMENTAL DETAILS

Two samples were studied experimentally. The first one consisted of a polished (111) single crystal of silicon on which a thermal oxide was grown. The second sample was prepared by depositing a thin film of poly(styrene) on top of a thermal oxide layer on a similar silicon substrate. The poly(styrene) was dissolved in toluene and deposited by spin-coating from the solution at 2000 rpm for 30 s.

Ellipsometry data were obtained on a rotating-analyzer spectroscopic ellipsometer VASE [1,40] as a function of both φ_0 and λ . Angular scans were performed in air over angles ranging from 30° to 85° in increments of 1° . Spectroscopic scans were typically performed over λ ranging from 300 to 800 nm in 10-nm increments. Typical standard deviations in the data were 0.04° for ψ and 0.1° for Δ .

The thicknesses of the same films were measured with Rutherford backscattering spectrometry (RBS) using the University of Surrey 2-MV Van de Graaff [41]. A 1.5-MeV $^4\text{He}^+$ beam at normal incidence was employed. The backscattered particles were detected at a 165° scattering angle in the same plane as the beam and the normal to the samples

(IBM geometry), and the detector resolution was 16-keV full width at half maximum.

VI. RESULTS AND DISCUSSION

A. Analysis of simulated test data

Before applying the algorithms described in the previous sections to real data, we tested them on simulated data from a predefined layer structure. This tactic has the advantage of eliminating uncertainties due to experimental errors and, more importantly, due to the limitations in the knowledge about any given real sample. By using simulated data we are certain of the original structure when comparing with the results obtained with SA and the MCMC algorithm. In simulated data there are no experimental errors, so we take the σ_i values from Eq. (21) to be equal, and require that on average 10% of all proposed transitions are accepted. The approach produced, in this case, an unrealistically high error in the determination of Δ and Ψ of about 5° and 1° , respectively, and was chosen to obtain a sufficiently long Markov chain in a reasonable calculation time.

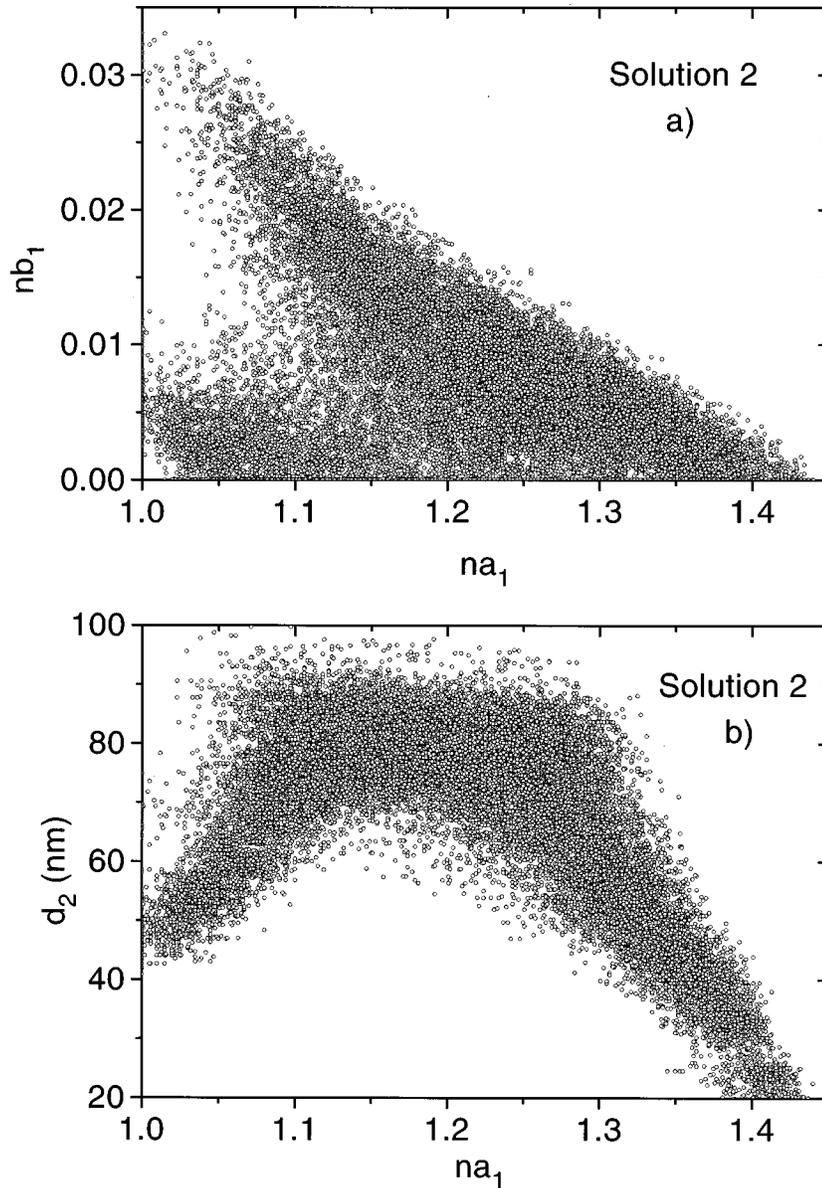


FIG. 5. Solution 2 from a MCMC analysis of simulated ellipsometry data (based on the structure described in Table I) shown as (a) nb_1 vs na_1 and (b) d_2 vs na_1 . Each dot in the figures represents one element x_i of the Markov chain.

For the test structure given in Table I we inserted values into Eq. (14) and substituted into Eq. (1) to generate the data shown in Fig. 1. Two angular scans with φ_0 varying between 30° and 85° were generated, for $\lambda = 500$ and 700 nm. Two spectroscopic scans were also generated, with λ ranging from 300 to 800 nm with $\varphi_0 = 65^\circ$ and 55° .

We then used an SA algorithm to fit these simulated data. We treated the thickness of each of the two layers (d_1 and

d_2) and two components of the refractive index (na_j and nb_j) of the two layers as free parameters, to make a total of six free parameters. While the fits obtained were nearly perfect, different solutions were obtained each time different random number sequences were used. This result means that even from such a simple structure (two transparent layers on a transparent substrate), and with a large amount of data available, the problem is still multimodal; each solution is

TABLE II. Averages and standard deviations (in parentheses) of the parameters for the MCMC solutions that reproduce the simulated ellipsometry data for a bilayer structure.

Solution	d_1 (nm)	na_1	nb_1	d_2 (nm)	na_2	nb_2
original structure	10.0	1.25	0.0015	55.0	1.45	0.0030
1	9.8(2.1)	1.15(8)	0.0080(65)	54.4(7.4)	1.44(1)	0.0036(13)
2a	17.7(3.4)	1.29(12)	0.0049(47)	49.5(18.5)	1.44(4)	0.0045(45)
2b	17.1(3.0)	1.07(4)	0.0002(4)	61.8(8.6)	1.45(2)	0.0021(14)
2c	16.5(2.2)	1.19(7)	0.0088(63)	79.5(5.3)	1.476(5)	0.0003(5)
3	50.2(8.7)	1.399(8)	0.0056(13)	2.2(2.0)	1.7(4)	0.026(14)

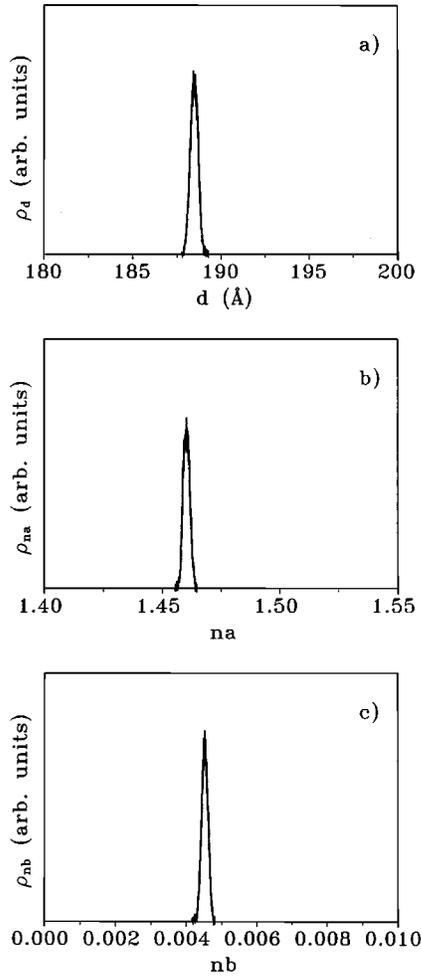


FIG. 6. Density of states obtained with a MCMC analysis for (a) d , (b) na , and (c) nb of a thin SiO_2 layer on a silicon substrate.

ambiguous and not necessarily the original structure used in the data simulation.

We next applied a MCMC algorithm to the same data, in order to investigate the whole parameter space. The calculated marginal densities describing the posterior distributions $p(\mathbf{x}|dI)$ are shown in Fig. 2. Here $p(\mathbf{x}|dI) \equiv (\rho_{d1}, \rho_{na1}, \rho_{nb1}, \rho_{d2}, \rho_{na2}, \rho_{nb2})$, where ρ_j reflects the density of states, or in other words, the density of solutions, that are consistent with the data. Figure 2 thereby represents the thickness and refractive index values that are consistent

TABLE III. Averages and standard deviations (in parentheses) of the thickness and refractive index parameters (na and nb) of a thin SiO_2 film on a Si substrate, as obtained with the MCMC algorithm and with least-squares fits using the literature value (Ref. [43]) for the bulk SiO_2 refractive index, letting the refractive index vary. The layer thickness determined with RBS was 18(2) nm.

Solution	d (nm)	na	nb
MCMC	18.85(2)	1.4602(14)	0.004 527(91)
Least-squares fit	19.25(3)	1.4476 ^a	0.003 666 ^a
Least-squares fit	18.85(13)	1.4601(82) ^b	0.004 532(59) ^b

^aLiterature values (Ref. [43]) for bulk SiO_2 , fixed in the fit.

^bAllowed to vary in the fit.

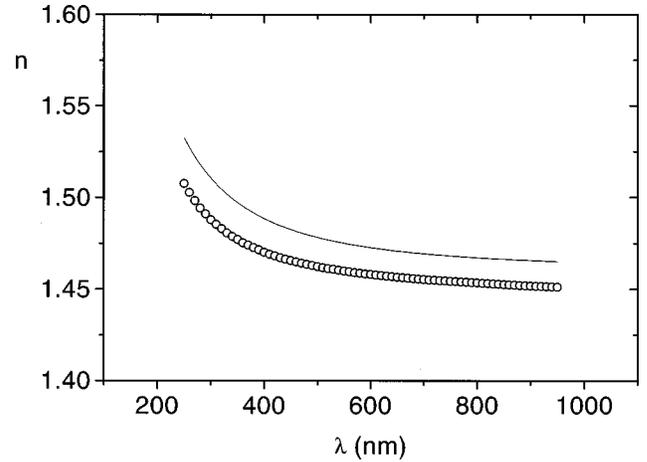


FIG. 7. Optical dispersion of the real part of the refractive index obtained from a MCMC analysis of ellipsometry data from a thin SiO_2 layer on a silicon substrate (shown as the solid line), compared to the literature values of refractive index for bulk SiO_2 (shown as dots).

with the data. The multimodality of the problem is clear. For instance, three distinct peaks in ρ_{d1} can be seen for the thickness d_1 , centered at about 10, 15, and 50 nm. We will henceforth refer to these as solutions 1, 2, and 3, respectively. There is structure in the other parameters as well. The thickness d_2 has a very strong peak at 0 nm, which corresponds to a single-layer solution, and also has two other broad peaks centered at about 50 and 80 nm. Both components of the refractive index of layer 1 (na_1 and nb_1) have a well-defined peak on top of a diffuse background. na_2 has two well-defined peaks on top of a diffuse background, while a single peak on top of a flat background can be distinguished in nb_2 .

In order to discover possible correlation between the different parameters, we calculated the partial density of states ρ_{na1} , ρ_{nb1} , ρ_{d2} , ρ_{na2} , and ρ_{nb2} corresponding to the three solutions 1, 2, and 3, defined as follows:

$$\text{solution 1 } d_1 < 13.5 \text{ nm,}$$

$$\text{solution 2 } 13.5 \text{ nm} < d_1 < 20 \text{ nm,}$$

$$\text{solution 3 } d_1 > 20 \text{ nm.} \quad (27)$$

The results obtained for all three solutions are shown in Fig. 3. It is clear that solution 3 is the single-layer solution, as the peak centered at $d_1 \approx 50$ nm is correlated to the strong peak at $d_2 \approx 0$ nm and to the corresponding undetermined refractive index for layer 2. It is also apparent that solution 1 corresponds to $d_2 \approx 50$ nm. Solution 2 seems to correspond to $d_2 \approx 80$ nm but also includes a tail that extends to low d_2 values. Finally, while the values of na_2 for solutions 1 and 2 and of nb_2 for solution 1 are well-defined peaks, we observe broad distributions in na_1 and nb_1 for solutions 1 and 2 and nb_2 for solution 2. These results are summarized in Table I.

The broad distributions of na_1 , nb_1 , nb_2 , and d_2 can hide additional structure. To uncover any hidden structure, in Fig. 4, we plotted, for solution 1, nb_1 as a function of na_1 and nb_2 as a function of na_1 . Likewise, in Fig. 5, for solu-

TABLE IV. Averages and standard deviations of the thickness and refractive index obtained from a poly(styrene)/SiO₂ bilayer on a Si substrate, as obtained by several means: a least-squares fit that uses the known SiO₂ refractive index, with and without allowing for a Δd_1 nonuniformity in the thickness of the poly(styrene) layer; and MCMC analysis, for the whole range of solutions found and restricted to the most probable solution.

Solution	d_1 (nm)	na_1	nb_1	d_2 (nm)	na_2	nb_2	Comments
bilayer of poly(styrene) and SiO ₂	36.9–45.8 ^a			93.4(1.8) ^b	1.4476 ^c	0.003 666 ^c	
least squares with nonuniformity	36.1(0.7)	1.557(3)	0.028(2)	97.7(0.1)	1.4476 ^d	0.003666 ^d	$\Delta d_1 = 4.45(0.37)$
least squares	37.2(0.8)	1.556(9)	0.024(6)	97.7(0.1)	1.4476 ^d	0.003666 ^d	$\Delta d_1 = 0$
MCMC	48.5(26.9)	1.52(5)	0.01(1)	90.5(25.7)	1.482(16)	0.0033(30)	$d_1 + d_2 = 139(2)$
MCMC restricted	28.4(6.7)	1.54(4)	0.016(11)	109.2(6.1)	1.483(11)	0.0031(22)	$10 < d_1 < 40$ $90 < d_2 < 120$

^aDetermined with RBS. The two values indicate spots in the edge and center of the sample.

^bDetermined with RBS.

^cSiO₂ values from the literature (Ref. [43]).

^dHeld constant during the least-squares fitting procedure.

tion 2 we plotted nb_1 as a function of na_1 and d_2 as a function of na_1 . Each point in these plots represents one individual solution obtained, that is, one element \mathbf{x}_i of the Markov chain. The figures can then be seen as phase diagrams where the layer structures (within each of the solutions) that are consistent with the data lie. No extra structure can be found in solution 1, except that nb_1 and na_1 are related to each other within a broadband: $-0.075(na_1 - 1.07) < nb_1 < -0.075(na_1 - 1.4)$. Solution 2 is more interesting. It is clear from Fig. 5(a) that it corresponds to at least two different solutions, as two more or less disjoint areas can be observed. Furthermore, a closer look at Fig. 5(b) reveals that three different solutions are present. One solution, with d_2 centered around 80 nm corresponds to the peak that can be seen in Fig. 5(a). On the other hand, the tail at low d_2 values seen in Fig. 5(a) corresponds to two different solutions, one with low na_1 values and the other with high na_1 values. These can be seen in Fig. 5(b) as the two branches extending to low d_2 values. Calculation of partial density functions reveals that the low na_1 value branch is correlated with the (low nb_1)/(low na_1) region in Fig. 5(a), and that the high na_1 value branch is correlated with the (high nb_1)/(high na_1) region.

We are finally left with five distinct solutions, all of which can adequately reproduce the test data shown in Fig. 1. We calculated for each one of these the average and standard deviation for all parameters, which are given in Table II. Solution 1 is, within the error bounds, the original layer structure used to calculate the test data analysed. As it is not possible to decide which of the solutions found is the ‘‘correct’’ one from the ellipsometry data only, extra information would be needed in a real experiment. It should be noted, however, that in this case such information must include the thickness of layer 1, since it is the only parameter that is clearly different between solution 1 and the other possible solutions, within one standard deviation. Alternatively, information on the refractive index of both layers would also

suffice (taking into account that realistic errors smaller than 1° would lead to smaller standard deviation values): the only solution with similar refractive index of the first layer is solution 2c, which can be distinguished from solution 1 with basis on the refractive index of the second layer.

Finally, the MCMC analysis tested only two-layer structures (although with the possibility of reducing their thickness to zero). Allowing the existence of extra layers would be almost certain to increase the number of possible solutions.

B. Analysis of a thin SiO₂ film on Si

The thickness of the thermal oxide on silicon substrate was determined to be 18(2) nm using RBS analysis, assuming the bulk density [42] of SiO₂ of 6.6×10^{22} at/cm³ in the

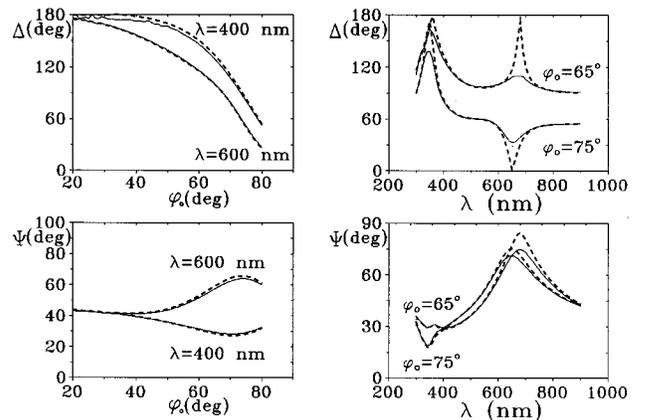


FIG. 8. VASE ellipsometry data measured from a poly(styrene) (layer 1)/SiO₂ (layer 2) bilayer on a silicon substrate in air (solid line). The dashed lines are the best-fit result from SA using a model of two homogeneous layers.

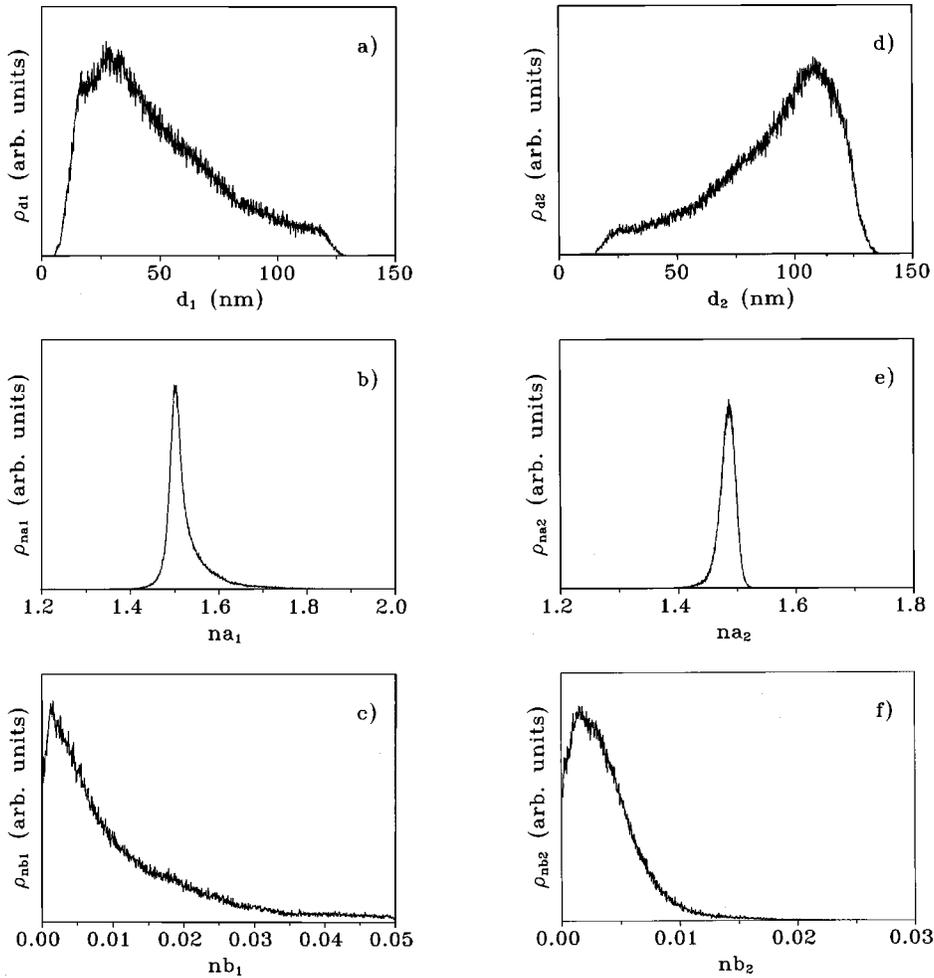


FIG. 9. Density of states obtained with a MCMC analysis of ellipsometry data from the poly(styrene)/SiO₂ bilayer on a silicon substrate showing results for (a) d_1 , (b) na_1 , (c) nb_1 , (d) d_2 , (e) na_2 , and (f) nb_2 .

data fitting. SA and MCMC analyses were performed on a set of ellipsometry data obtained from three spectroscopic scans ($300 \text{ nm} < \lambda < 700 \text{ nm}$) performed at $\varphi_0 = 72^\circ$, 75° , and 78° . The experimental errors were included in the analysis as given in Eq. (21), and hence the confidence limits in the solutions calculated with the MCMC algorithm reflect the exact error structure of the problem. The posterior distribu-

tion $p(\mathbf{x}|dI) \equiv (\rho_d, \rho_{na}, \rho_{nb})$ is shown in Fig. 6. A single very well-defined solution has been obtained, which means that the problem is fully unambiguous. This result is expected in this trivial case, since the problem could also be solved by exact data inversion. Nevertheless, obtaining an expected result provides confidence in the SA-MCMC method.

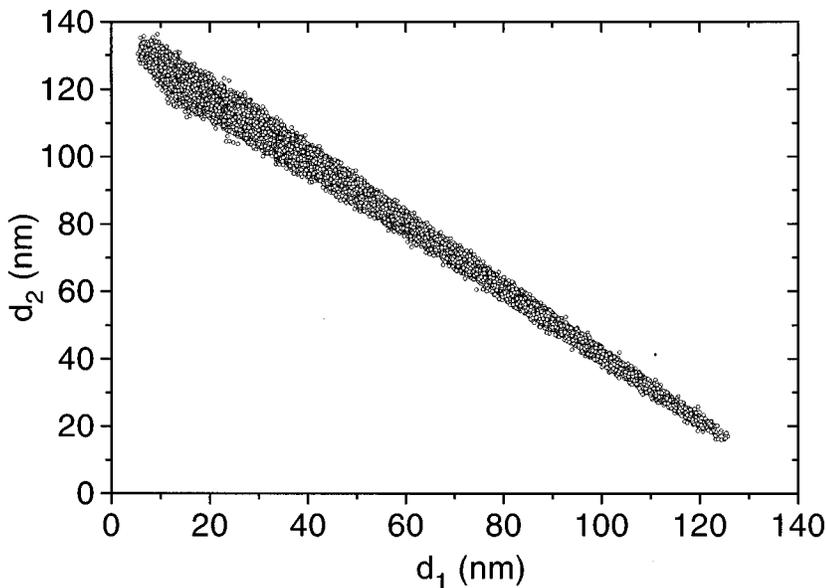


FIG. 10. Results of a MCMC analysis for the poly(styrene)/SiO₂ bilayer on a silicon substrate showing solutions as d_2 vs d_1 .

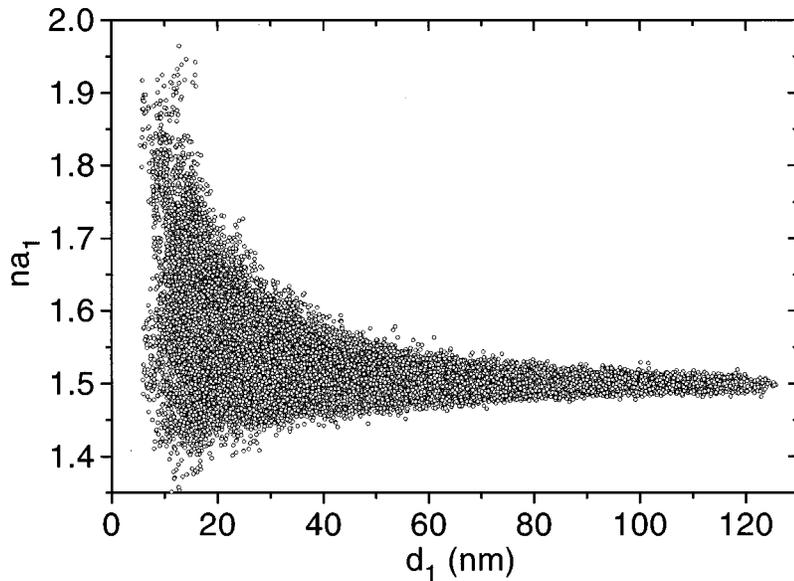


FIG. 11. Results of a MCMC analysis for the poly(styrene)/SiO₂ bilayer on a silicon substrate showing solutions as na_1 vs d_1 .

The average and standard deviation of the thickness and refractive index of the layer are given in Table III. The table also reports the SiO₂ thickness and refractive index parameters (na and nb) that were obtained from a Levenberg-Marquardt least-squares fit to the same data performed with commercial software [40] using the literature values [43] for the refractive indices of bulk Si and SiO₂. The refractive indices obtained from MCMC analysis is slightly higher than the literature values for the bulk material, but the same optical dispersion as found in bulk SiO₂ is obtained, as shown in Fig. 7. The difference in index could reflect real differences in structure and density between the thin film and the bulk material. If the refractive index parameters are allowed to vary in the least-squares fit, then very similar values are obtained as from the MCMC analysis, as seen in Table III. Finally, when analysing a single spectroscopic scan (at $\varphi_0 = 72^\circ$), the same single well-defined solution is obtained, which means that the scans at $\varphi_0 = 72^\circ$, 75° , and 78° contain redundant information.

C. Analysis of a poly(styrene)/SiO₂ bilayer on Si

The thicknesses of each of the two layers in the poly(styrene)/SiO₂ bilayer were determined with RBS, assuming the bulk density of polystyrene [44] of 8.4×10^{22} at/cm³ and of SiO₂ as before. The thickness of the oxide layer was found to be 93.4 nm. The thickness of the poly(styrene) layer was measured at two positions on the sample. Two different values (36.9 and 45.8 nm) were obtained, which indicates that the layer thickness is inhomogeneous, as is sometimes found in films deposited by spin coating. Table IV summarizes these results from RBS. Ellipsometry data, from both angular and spectroscopic scans, were obtained from the same bilayer sample. These data are shown in Fig. 8.

We first obtained a least-squares fit to the ellipsometry data using commercial software [40] that makes use of the Levenberg-Marquardt algorithm. The resulting parameters are given in Table IV. The fits, however, are not good, as can be seen in Fig. 8. A good fit could only be obtained by

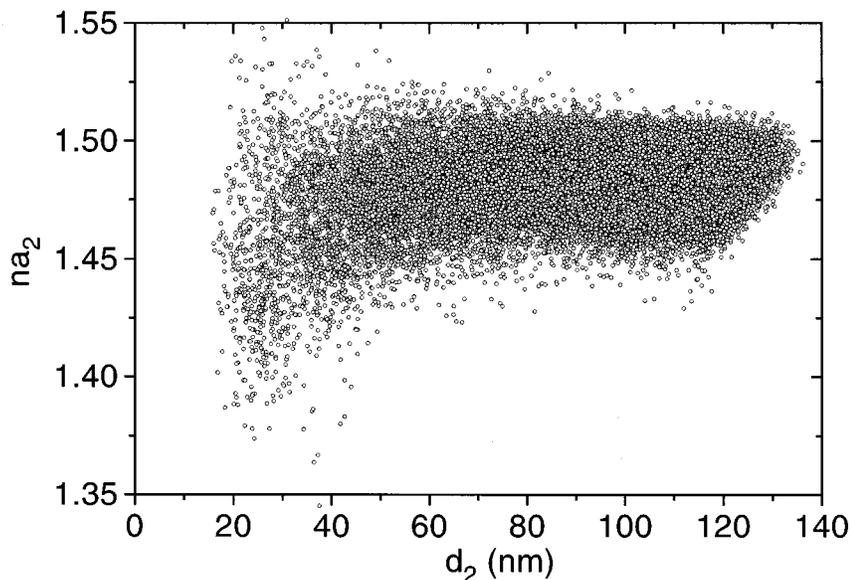


FIG. 12. Results of a MCMC analysis for the poly(styrene)/SiO₂ bilayer on a silicon substrate showing solutions as na_2 vs d_2 .

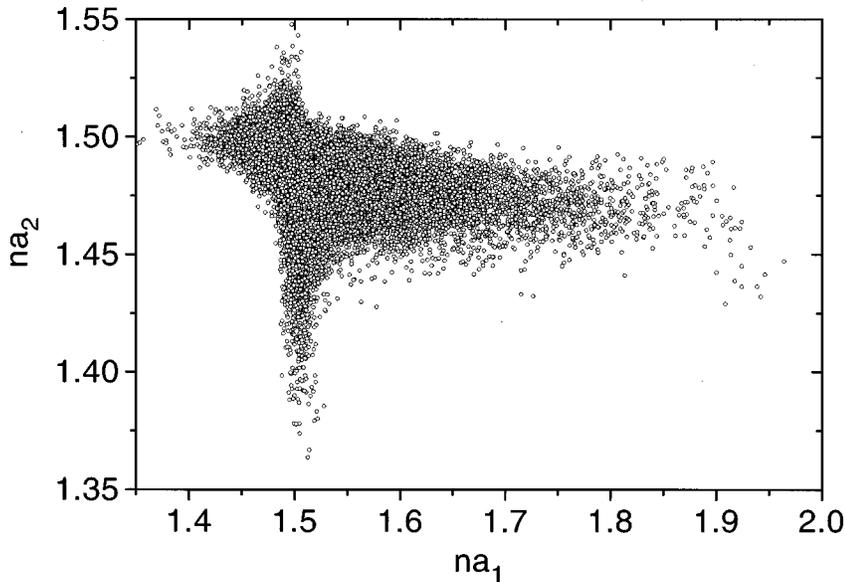


FIG. 13. Results of a MCMC analysis for the poly(styrene)/SiO₂ bilayer on a silicon substrate showing solutions as na_2 vs na_1 .

assuming a model in which thickness non-uniformity of the poly(styrene) layer is allowed. The results from this refined fitting model are listed in Table IV. However, in the model that we implemented, no thickness nonuniformity is considered, which means that a perfect and realistic fit to the data can never be obtained. As the χ^2 would be due to a limitation in the model and not to experimental error, the results obtained would not reflect the structure of the problem but limitations of the model instead. In order to overcome this problem, we took as the error bar for each data point, not the experimental error, but the deviation between the data and the best-fit assuming homogeneous layers. The minimum error in each point was set to 1° . This approach is *ad hoc* to the extent that only by taking into account all physical effects, including thickness nonuniformity, would one obtain absolutely accurate marginal densities. This method ensures however, that all solutions similar to the best fit will be highly probable in the MCMC calculation, and therefore will reflect the real structure of the problem.

The results obtained for the poly(styrene)/SiO₂ bilayers are shown in Fig. 9. The thickness of both layers is clearly ambiguous, as d_1 takes values between 10 and 120 nm, and d_2 falls between 20 and 130 nm. Note that a single layer model (with d_1 or $d_2=0$) is not an acceptable solution. On the other hand, the na_j values for the two layers are well defined. The wavelength-dependent components have a strong peak at low nb_j values and a tail extending to high values. The tail is more pronounced for the poly(styrene) layer. It should be noted that, although close, the na_j values are different in the two layers: slightly above and slightly below 1.5 for the poly(styrene) (layer 1) and SiO₂ (layer 2), respectively.

Figure 10 shows that the thickness of the two layers are strongly linearly correlated: $d_1 + d_2 \approx 140$ nm. This result probably stems from their similar refractive index values and the associated weak reflection from the interface between the two layers. The order of the layers are not, however, interchangeable: it is clear that the SiO₂ film is *not* on top of the poly(styrene) layer. Moreover, no single-layer solution corresponding to an average refractive index value is obtained. The behavior of the refractive index values in the two layers

is quite different, as can be seen in Figs. 11 and 12. For the most probable value of d_1 [i.e., the poly(styrene) layer], which is in the range between 10 and 40 nm the value of na_1 is poorly defined, with a large scattering of values apparent. For the SiO₂ layer, in contrast, the most probable d_2 values (around 90–120 nm) are related to a well-defined range of na_2 values. The independence between the refractive index values of the two layers can also be seen by plotting na_1 vs na_2 , which is done in Fig. 13. For $na_1 \approx 1.5$, na_2 can take any value within its allowed range; and for $na_2 \approx 1.48$, na_1 can take any value within its allowed range.

Finally, the average and standard deviation of the thickness and refractive index components of the two layers, as obtained with MCMC analysis, are given in Table IV. The results obtained for the most probable solution ($d_1 \approx 10$ –40 nm and $d_2 \approx 90$ –120 nm) are also given, and they match well the best fit obtained.

VII. CONCLUDING REMARKS

As far as we are aware, the SA-MCMC approach demonstrated here is the first systematic method that can analyze VASE data without any knowledge of the thin film structure, and provide a *guarantee* that the global minima in the error will always be reached. The method finds *all* possible structures that can produce an ellipsometry data set and calculates the errors on the fits. Although other methods could be developed to search the solution space in a systematic fashion, these are unlikely to be guaranteed to find all solutions. The capabilities of the SA-MCMC method increase the applications of ellipsometry to include the analysis of unknown multilayer samples. The SA-MCMC method is therefore an intriguing and attractive alternative to the Levenberg-Marquardt and similar least-squares methods often used to analyze VASE data. Our analysis of these relatively simple structures can be extended to more complicated systems consisting of three or more layers, having an unknown substrate or ambient, and having surface roughness and biaxiality. The computations shown in this paper have, as yet, not been optimized for speed, but we expect that significant improvements are possible.

ACKNOWLEDGMENTS

N.P.B. was partially supported by the U.K. Engineering and Physical Sciences Research Council (EPSRC) through Grant No. GR/L78512. We also thank the EPSRC for a grant

(Grant No. GR/L12066) toward the purchase of the ellipsometer used in this work and for financial support for R.S. Finally, the comments of Dr. Chris Jeynes on a first draft of the paper are greatly appreciated.

-
- [1] *Polymer Surfaces and Interfaces III*, edited by R. W. Richards and S. K. Peace (Wiley, Chichester, 1999), Chap. 1.
- [2] E. H. Korte and A. Röseler, *Analyst* (Cambridge, U.K.) **123**, 647 (1998).
- [3] J. L. Keddle, R. A. L. Jones, and R. A. Cory, *Europhys. Lett.* **27**, 59 (1994).
- [4] J. S. Papanu, D. W. Hess, A. T. Bell, and D. S. Soane, *J. Electrochem. Soc.* **136**, 1195 (1989).
- [5] M. Malmsten, *J. Colloid. Interface Sci.* **166**, 333 (1994).
- [6] J. A. De Feijter, J. Benjamins, and F. A. Veer, *Biopolymers* **17**, 1759 (1978).
- [7] A. Roseler, R. Dietel, and E. H. Korte, *Mikrochim. Acta* **S14**, 657 (1997).
- [8] A. Markwitz, H. Baumann, W. Grill, B. Heinz, A. Roseler, E. F. Krimmel, and K. Bethge, *Fresenius J. Anal. Chem.* **353**, 734 (1995).
- [9] R. M. A. Azzam and N. M. Bashara, *Ellipsometry and Polarized Light* (North-Holland, Amsterdam, 1977), Chap. 4.
- [10] J.-P. Drolet, S. C. Russev, M. I. Boyanov, and R. M. Leblanc, *J. Opt. Soc. Am. A* **11**, 3284 (1994).
- [11] B. D. Johs, W. A. McGahan, and J. A. Woollam, *Thin Solid Films* **253**, 25 (1994).
- [12] M. Born and E. Wolf, *Principles of Optics* (Cambridge University Press, Cambridge, 1997), pp. 95 and 96.
- [13] A. R. Reinberg, *Appl. Opt.* **11**, 1273 (1972).
- [14] G. H. Park, Y. H. Pao, B. Igel'nik, K. G. Eyink, and S. R. Leclair, *IEEE Trans. Neural Netw.* **7**, 816 (1996).
- [15] D. W. Marquardt, *J. Soc. Ind. Appl. Math.* **11**, 431 (1963).
- [16] A. Zuber, H. Jänchen, and N. Kaiser, *Appl. Opt.* **35**, 5553 (1996).
- [17] M. Fried and P. Masa, *J. Appl. Phys.* **75**, 2194 (1994).
- [18] J. C. Charmet and P. G. de Gennes, *J. Opt. Soc. Am.* **73**, 1777 (1983).
- [19] V. G. Polovinkin and S. N. Svitashva, *Thin Solid Films* **313-314**, 128 (1998).
- [20] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, *Science* **220**, 671 (1983).
- [21] Emile Aarts and Jan Korst, *Simulated Annealing and Boltzmann Machines: A Stochastic Approach to Combinatorial Optimization and Neural Computing* (Wiley, Chichester, 1989).
- [22] R. H. J. M. Otten and L. P. P. van Ginneken, *The Annealing Algorithm* (Kluwer, Norwell, MA, 1989).
- [23] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran*, 2nd ed. (Cambridge University Press, Cambridge, 1992), p. 438.
- [24] N. P. Barradas, C. Jeynes, and R. Webb, *Appl. Phys. Lett.* **71**, 291 (1997).
- [25] N. P. Barradas, P. K. Marriott, C. Jeynes, and R. P. Webb, *Nucl. Instrum. Methods Phys. Res. B* **136-138**, 1157 (1998).
- [26] N. P. Barradas, C. Jeynes, R. P. Webb, U. Kreissig, and R. Grötzschel, *Nucl. Instrum. Methods Phys. Res. B* **149**, 233 (1999).
- [27] N. P. Barradas, C. Jeynes, and M. A. Harry, *Nucl. Instrum. Methods Phys. Res. B* **136-138**, 1163 (1998).
- [28] N. P. Barradas, C. Jeynes, and S. M. Jackson, *Nucl. Instrum. Methods Phys. Res. B* **136-138**, 1168 (1998).
- [29] N. P. Barradas, C. Jeynes, K. P. Homewood, B. J. Sealy, and M. Milosavljevic, *Nucl. Instrum. Methods Phys. Res. B* **139**, 235 (1998).
- [30] Y. A. Wilks, B. M. Slator, and L. M. Guthrie, *Electric Words: Dictionaries, Computers, and Meanings* (MIT Press, Cambridge, 1996), p. 203–206.
- [31] *Maximum Entropy and Bayesian Methods in Inverse Problems*, edited by C. R. Smith and W. T. Grandy (Reidel, Dordrecht, 1985).
- [32] *Maximum Entropy in Action*, edited by B. Buck and V. A. Macaulay (Clarendon, Oxford, 1994).
- [33] P. M. Lee, *Bayesian Statistics: An Introduction*, 2nd ed. (Arnold, London, 1997).
- [34] R. Fischer, M. Mayer, W. von der Linden, and V. Dose, *Phys. Rev. E* **55**, 1 (1997).
- [35] W. von der Linden, M. Donath, and V. Dose, *Phys. Rev. Lett.* **71**, 899 (1993).
- [36] N. P. Barradas, A. P. Knights, C. Jeynes, O. A. Mironov, T. J. Grasby, and E. H. C. Parker, *Phys. Rev. B* **59**, 5097 (1999).
- [37] N. P. Barradas, C. Jeynes, M. Jenkin, and P. K. Marriott, *Thin Solid Films* (to be published).
- [38] N. Metropolis, A. W. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- [39] P. R. Bevington and D. K. Robinson, *Data Reduction and Error Analysis for the Physical Sciences*, 2nd ed. (McGraw-Hill, New York, 1994), pp. 65–72.
- [40] WVASE32 Software, Version 3.148, J. A. Woollam Co., Inc., Lincoln, NE.
- [41] C. Jeynes, N. P. Barradas, M. J. Blewett, and R. P. Webb, *Nucl. Instrum. Methods Phys. Res. B* **136-138**, 1229 (1998).
- [42] *CRC Handbook of Chemistry and Physics*, 75th ed. edited by D. R. Lide and H. P. R. Frederikse (CRC Press, Boca Raton, FL, 1995), pp. 4–148.
- [43] *Handbook of Optical Constants*, edited by E. D. Palik (Academic, London, 1991), pp. 547 and 749.
- [44] *Polymer Handbook*, 3rd ed. edited by J. Brandrup and E. H. Immergut (Wiley, Chichester, 1989), p. V/82.