View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by HAL-CEA

# An inverse ellipsometric problem for thin film characterization: comparison of different optimization methods

Ayse Akbalik, Sébastien Soulan, Jean-Hervé Tortai, David Fuard, Issiaka Kone, Jérôme Hazart, Patrick Schiavone

▶ To cite this version:

Ayse Akbalik, Sébastien Soulan, Jean-Hervé Tortai, David Fuard, Issiaka Kone, et al.. An inverse ellipsometric problem for thin film characterization: comparison of different optimization methods. John A. Allgair, Christopher J. Raymond. Metrology, Inspection, and Process Control for Microlithography XXIII, Feb 2009, San Jose (CA), United States. SPIE, 7272, in press, 2009, <10.1117/12.814118>. <hal-00374509>

**HAL Id: hal-00374509**

**https://hal.archives-ouvertes.fr/hal-00374509**

Submitted on 8 Apr 2009

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# An inverse ellipsometric problem for thin film characterization: Comparison of different optimization methods

Ayşe Akbalık[a], Sébastien Soulan[a], Jean-Hervé Tortai[a], David Fuard[a], Issiaka Kone[a],Jérôme Hazart[b], Patrick Schiavone[a,c]

[a]Laboratoire des Technologies de la Microélectronique CNRS, c/o CEA, 17 av. des martyrs, 38054 Grenoble Cedex, France

[b] LETI, CEA, 17 av. des martyrs, 38054 Grenoble Cedex, France

[c] UMI2958 CNRS-Georgia Tech, Georgia Institute of Technology, Atlanta, GA 30332

**ABSTRACT**

In this paper, an ill-posed inverse ellipsometric problem for thin film characterization is studied. The aim is to determine the thickness, the refractive index and the coefficient of extinction of homogeneous films deposited on a substrate without assuming any *a priori* knowledge of the dispersion law. Different methods are implemented for the benchmark. The first method considers the spectroscopic ellipsometer as an addition of single wavelength ellipsometers coupled only via the film thickness. The second is an improvement of the first one and uses Tikhonov regularization in order to smooth out the parameter curve. Cross-validation technique is used to determine the best regularization coefficient. The third method consists in a library searching. The aim is to choose the best combination of parameters inside a pre-computed library. In order to be more accurate, we also used multi-angle and multi-thickness measurements combined with the Tikhonov regularization method. This complementary approach is also part of the benchmark. The same polymer resist material is used as the thin film under test, with two different thicknesses and three angles of measurement. The paper discloses the results obtained with these different methods and provides elements for the choice of the most efficient strategy.

**Keywords:** Inverse ellipsometric problem, Tikhonov regularization, cross validation, library search

## 1. INTRODUCTION

Ellipsometry is a very commonly used optical technique that proved to be highly accurate for numerous applications such as thin film thickness measurement. It measures the change in the polarization state of the incident light reflected from the surface of a sample. The ellipsometry uses the properties of the reflected light, in order to search for some physical or optical properties of the sample. Reader can refer for example to the complete reference on ellipsometry edited by Tompkins and Irene [1] for further details.

However, especially when dealing with unknown materials or absorbing films, the reliability of the ellipsometric measurement depends on the accurate determination of the refractive index of the material under test. In the case of spectroscopic ellipsometry, the complex refractive index needs to be known on the full wavelength range. This is generally done using various dispersion functions (Cauchy, oscillators, ...) that describe the variation of the optical constants of a material versus wavelength. This brings the user to a severe issue which is the choice of the dispersion curve to be used for his material. This has to be done *a priori* and very often these dispersion functions rely on strong hypothesis on the physics of the material. In order to overcome that issue, we tried to develop strategies for the determination of the refractive index that does not rely on any *a priori* hypothesis, except regularity. The present paper aims at describing several of these strategies, most of them being original, and compare their output when applied to the same polymer material.

In the study of the inverse ellipsometric problem for thin film characterization, the associated numerical problem is ill-conditioned. The three conditions given by Hadamard (existence, uniqueness and stability of the solutions) are not met. Different solutions can give the same resulting curves and due to noisy measurements, the solutions are not always stable. To solve it, we use several methods found in the inverse problem literature. Before describing the problem in detail, we give a brief description of the ellipsometry and an overview of the inverse ellipsometric problem literature.

Our aim is to determine the thickness ($e$), the refractive index ($n_i$) and the coefficient of extinction ($k_i$) of homogeneous films deposited on a Silicon substrate. In this problem, $n_i$ and $k_i$ are both dependent on the wavelength (or energy-eV). When $N$ different wavelengths are used, there are thus $2N+1$ unknowns. The data available to make this characterization are $2N$ measured values provided by the spectroscopic ellipsometer, denoted by $Is_i^m$ and $Ic_i^m$ , for $1< i < N$. The physical meanings of $Is$ and $Ic$ vectors are explained in the next section. We will see in Section 3 that it is also theoretically possible to compute these $2N$ values of $Is$ and $Ic$ using the unknown parameters ($e$, $n_i$, $k_i$). One should thus estimate the values of these parameters in order to fit best the measured data. The best parameters are those matching the best the measured data $Is^m$ and $Ic^m$. In order to determine these $2N+1$ unknowns, we use three different methods and compare the results obtained by each one.

In the literature, one can find many papers studying inverse problems. Without being very exhaustive, we give an overview of the papers which are the most related with our study. There are many studies on the methods used to solve the ill-conditioned inverse problems. For a very detailed reference on inverse problems, reader can be refered to Tarantola [2], where the author presents discrete and general inverse problem theory. Haber [3] deals with the numerical solutions of linear and nonlinear inverse problems and develops new techniques for solving such problems. Generalized Cross Validation and Krylov space methods are used respectively for noise estimation and large system inversion for linear systems. A new algorithm is developed for the non-linear case, based on the differentiation between the correlated errors coming from linearization and non-correlated noise coming from measurements. Another interesting study on ill-posed problems is from Hansen [4]. The author gives numerical methods for solving rank-deficient problems and linear discrete ill-posed problems. New and existing numerical algorithms are studied. Regularization methods and parameter choice methods are detailed. Busby and Trujillo [5] study two methods in order to choose the best regularization parameter for the Tikhonov regularization method: L-curve and generalized cross validation (GCV). They indicate that the two methods give for most cases the same parameter. L-curve needs less computing, but works well for linear systems only. On the contrary, GCV takes additional computation resources but it can be applied to non-linear systems. Krawczyk-Stando and Rudnicki [6] compare also two methods to choose the best regularization parameter, issued from the standard Tikhonov regularization method: the well-known L-curve and the U-curve method which is a new criterion that authors propose. These two methods can be used for linear systems, which is not the case for our problem. On the examples they test, the U-Curve is found to have a smaller error than L-curve.
We give some papers dealing with inverse problems in ellipsometry. Bobro *et al*. [7] study an ill-posed inverse ellipsometric problem and try to obtain more stable solutions than given in the literature. They use some new criteria to choose the optimal parameters for ultra thin films. They solve a multidimensional problem where multi-angle measurements are used as data. They run many numerical simulations and use optimization methods from the literature (i.e. Box's complex method). However, with this method they are limited to only 10 decision variables, which are the thickness and the refractive index dependent on the angle of measurement. Nevertheless they announce that they can successfully study surface films with thickness ranged from 2 to 10 nm. Rosa [8] studies also the inverse problem of ellipsometry for thin films in order to determine the thickness, the refractive index and the coefficient of extinction. The author uses a statistical method, called Bootstrap, in order to determine the solution and also the standard errors associated to it with an error propagation. Two versions of the Bootstrap method are used: using an initial data set or generating them artificially by the Monte Carlo method. An advantage of this method is indicated as the accuracy assigned to the final results. The author emphasizes the efficiency and the ease of use of this method not only in ellipsometric problems but also in the general field of inverse problems. Tonova and Konova [9] propose an algorithm to obtain the refractive index and the thickness of inhomogeneous films. They use multiple angle of incidence data. In their previous works, the same authors developed algorithms to find the refractive index with a known thickness. In this paper they extend this study and give a so-called Newton-Kantorovitch algorithm to determine both parameters. This method is said to be analogous to the discrete Newton method for solving non-linear equations. They test it using simulated data for different inhomogeneous transparent films and show its efficiency. Prior to our study in this paper, Kone [10] studied the same problem with Tikhonov regularization approach, but using the L-curve method to find the best regularization parameter. As the problem is non-linear, the L-curve did not give the desired results. We extend the same problem proposing different optimization techniques for the non-linear case.

In the current paper, we present different techniques to solve the non-linear, ill-posed inverse problem for the ellipsometry. We begin by presenting the problem and its theoretical formulation in Section 2. We detail the physical equations which allow to compute the $Is$ and $Ic$ curves theoretically, explaining the unknowns of the problem. In Section 3, we describe the methods we used, namely local optimization, Tikhonov regularization and library search. We provide computational results in Section 4 and finally some concluding remarks are presented in Section 5.

## 2. PROBLEM DESCRIPTION

The ellipsometer provides us with the measured data $Is_i^m$ and $Ic_i^m$, for the interval [1, ..., *N*] of wavelength (or energy-eV), where *i* represents the wavelength ( $i \in \mathbb{N}$ ) and *m* indicates that these are measured values. These two curves *Is* and *Ic* can also be computed theoretically, that will be denoted by $Is_i^t$ and $Ic_i^t$.

$I_s^t = \sin 2\psi \sin \Delta$, $I_c^t = \sin 2\psi \cos \Delta$ , for a 0 angle between the polarizer and the modulator (1)

$I_s^t = \sin 2\psi \sin \Delta$, $I_c^t = \cos 2\psi$ , for an angle π/4 between the polarizer and the modulator. (2)

In these two equations, $\psi$ and $\Delta$ are computed as follows:

$$\rho = \frac{r_p}{r_s} = \tan \psi \, e^{j\Delta} \quad (3)$$

Equation (3) represents the formalism which allows to compute $\rho$, the ratio of complex reflectivities. $r_p$ and $r_s$ are the Fresnel reflection coefficients for the p and s polarization respectively. They are a function of incidence angle, refractive angle, the refractive index ($n_i$), the coefficient of extinction of the sample ($k_i$), the thickness of the sample *e*, the wavelength $\lambda_i$, etc. tan $\psi$ is the attenuation ratio and Δ the phase shift between the parallel and perpendicular components of the polarized light. Once $\psi$ and Δ are known, one can theoretically compute $Is_i^t$ and $Ic_i^t$ using Equations (1) and (2), depending on the configuration of the ellipsometer.
One can easily see that, in order to compute $\rho$, the unknowns are the thickness *e*, the refractive index ($n_i$) and the coefficient of extinction of the sample ($k_i$). The other parameters are all known during the experience. We also know that we have only 2*N* given data versus 2*N*+1 unknown parameters. To avoid this difficulty, there exist different ways to increase the number of independent measurements. One can use multiple angle of incidence with the same sample, many measurements can be made for different thicknesses, etc. We used these two ways in this paper for the benchmark with the other methods. In the next section we detail the methods used.

## 3. DESCRIPTION OF THE METHODS

Before giving a brief description of each method, let us present several notations used in the following.

- $x^* = (e, n_i, k_i)$ : optimal vector to be determined, of dimension 2*N*+1
- *F*(*x*): function which computes theoretically $Is^t$ and $Ic^t$ for a given *x* (from Fresnel equations)
- *b*: measured values ($Is^m$ and $Ic^m$ measured by the ellipsometer)
- $\beta$: coefficient of regularization
- $\mathcal{L}$: regularization matrix (second derivative in the current case)

The spectroscopic ellipsometry problem of deducing the unknown parameters from a best match between the theoretical and measured response can be stated in the following way

$$x^* = argmin_x \{ \|F(x) - b\|_2^2 + \beta^2 \|\mathcal{L} x\|_2^2 \}$$

First, a simplification *a priori* can be done on the number of unknowns. For large wavelengths (small energy ranges) most of the thin films such as photo-resists or dielectrics that we are dealing with are known to be transparent. They are assumed to have a zero coefficient of extinction. For example photo-resists are dielectric materials that begin to absorb typically in the UV range. Hence in the infrared-visible range, no absorption is expected and $k_i$ is assumed to be 0. An initial estimation of the thickness can thus be done with only a subset of the 2*N* measurements. Secondly, the optical index of the film is assumed to follow a Cauchy model:
$n = A + 10^4 B/\lambda^2 + 10^9 C/\lambda^4$.

Hence inside the range of low energy (1.5eV, 4eV) an estimation of the thickness and of *n* can be made with only 3+1 unknowns. Once the thickness is known, different methods presented in this section allow us determining the *N* unknowns $n_i$ and the *N* unknowns $k_i$ that fit the best the experimental curves.

### 3.1 Optimization without regularization

The first method (sometimes denoted LO in the following) used in this paper consists in using an algorithm that will determine the $N$ $n_i$ and the $N$ $k_i$ that fit the best the experimental data wavelength by wavelength. It simply considers the spectroscopic ellipsometer as an addition of single wavelength ellipsometers coupled only via the film thickness. All minimization steps of the program use the lsqcurvefit routine implemented in Matlab. It minimizes in the least square sense, the errors between the measured $Is^m$ and $Ic^m$ and the $Is^t$ and $Ic^t$ theoretically computed. In our case, this routine uses a large scale algorithm based on the interior-reflective Newton method [11-12]. The solution is approximated using the method of preconditioned conjugate gradients.
The difficulty of this method is its dependence on the quality of the initial values of the parameters. It requires a good guess of the initial vector for the unknowns as a bad choice of this vector can result in an unacceptable physical solution due to local minima. This is a common difficulty for any non-linear minimization problem solved with local optimizers and is not easy to avoid. Thus, the user should pay attention to generate a correct initial vector. In order to solve this issue, a specific initial condition optimizer was computed.
Another disadvantage of this method is that the measurement noise, inherent to any experimental setup can significantly alter the stability of the solution and pass on the optimized parameters. We thus use another technique, called regularization, in order to smooth out the noisy parameters (see Section 3.2).

### 3.2 Tikhonov regularization

The second method is an improvement of the first one, using a technique called Tikhonov regularization. A second component weighted by a so called regularization coefficient $\beta$ is added in the objective function. It aims at providing a smoother solution. For more details on this technique see Tikhonov and Arsenine [13]. For instance, a very large $\beta$ may give a very smoothed curve since the weight of the regularization term is increased, however a very small $\beta$ can leave a noisy curve. $\beta$=0 means no regularity constraints, in this case the objective function is simply the sum of squared errors (i.e. the method of section 3.1). The important step is the determination of the best $\beta$ which provides the best balance between proximity to data and regularity of the solution. For that purpose, we use a cross-validation technique which is a measurement of the robustness of the computed model. At each step, one measured data is deleted from the data set and the optimization is done with the remaining data. The missing point is computed theoretically with an interpolation (or extrapolation). The error between the deleted measured value and the theoretically computed value is found and a sum is done for all missing points. The best $\beta$ is the one that minimizes the sum of errors in the least square sense.

$$\beta^{*}=argmin_{\beta}\{\sum_{i=1}^{N}(Is_i^t-Is_i^m)^2+(Ic_i^t-Ic_i^m)^2\}$$

To be more robust, we also use multi-angle and multi-thickness measurements together with Tikhonov regularization. Having more independent measurements, the resulting curves are found to be much closer to the physical values. Reader can find test results in Section 4.
A disadvantage of the cross-validation technique is the time needed for the computation of the best $\beta$. For instance, in our experiments, we use 101 wavelengths, resulting in 203 unknowns. For each $\beta$ value tested (dependent on the problem studied the total number of $\beta$ to test can reach 100), 101 measured values have to be deleted one by one to be able to compute the sum of errors. Thus, in total there are 10000 computations. Each one being an optimization problem with 100 unknown parameters.

### 3.3 Library Search

The third method consists of a library search. This is an application to thin film ellipsometry of a method developed in our group for dynamic scatterometry [14]. A highly dense library of $Is^t$ and $Ic^t$ values is generated for different combinations of $n$ and $k$. For a measured couple ($Is^m$,$Ic^m$), the aim is to choose the element in the library which minimizes a weighted sum of
- the proximity with the measured values of ($Is^m$,$Ic^m$)
- the regularity of the ($n$,$k$) variation through the wavelengths
As in the other methods, the thickness is assumed to be known for this technique. See Soulan [15] for more details on this library search method. One advantage of this technique is its speed to find the best parameter values. The generation of the library can be very time consuming, but once this is done, the optimization problem reduces to a database search.

In the next section, we give some results to compare all these methods.

# 4. EXPERIMENTAL RESULTS

The samples under test are two thin layers of the same resist. All data are measured on an UVISEL FUV AGAS from Horiba Jobin Yvon, spectroscopic ellipsometer equipped with rotating arms. The energy range used is 1.5eV to 6.5eV. Using the Cauchy approach described in Section 3, the thickness of each layer is deduced from the low energy part of the spectrum where the resist material is known to be transparent. The extracted thicknesses are 77.3nm and 103.1nm respectively. In this thickness range, no significant interface effect or difference in the deposition process is expected to induce variation of the complex refractive index which is therefore supposed to be identical for the two films. We performed multi- angle measurements : 3 different angles at 65°, 70° and 75° were used. This is aimed at improving the accuracy of the extracted optical parameters by mean of the use of a larger number of input data while keeping the same number of output parameters being optimized. First method, Tikhonov single angle method and library research method were run with experimental data collected with one incident angle of 70°.
The results of the three different approaches are summarized in Figure 1 and 2 for the thick and thin film respectively.

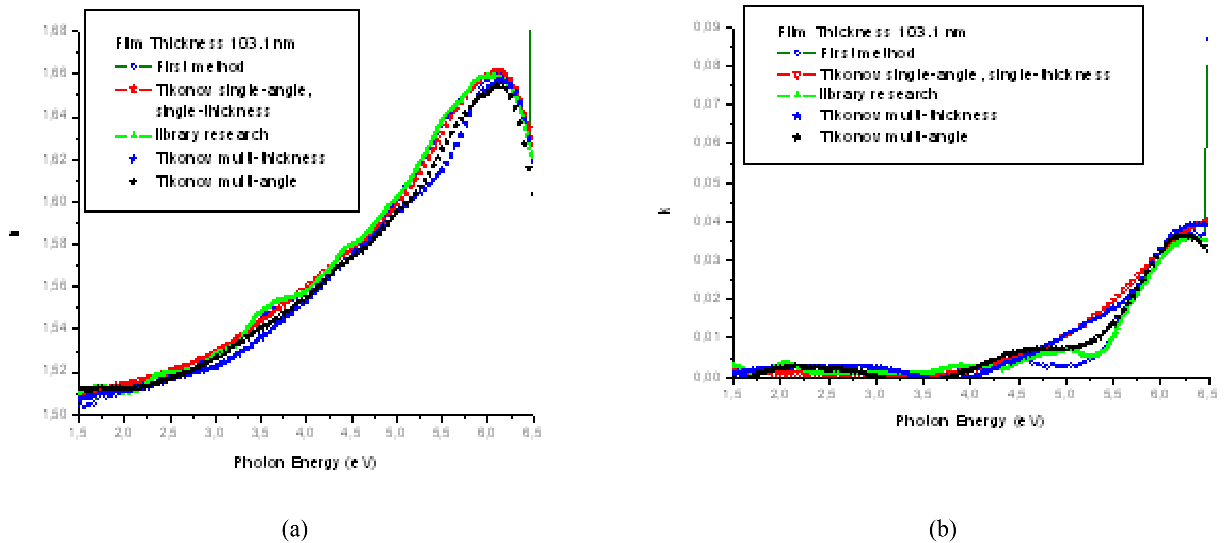


Figure 1: (a) *n* vs. photon energy and (b) *k* vs. photon energy (b) for the 103.1 nm thick material

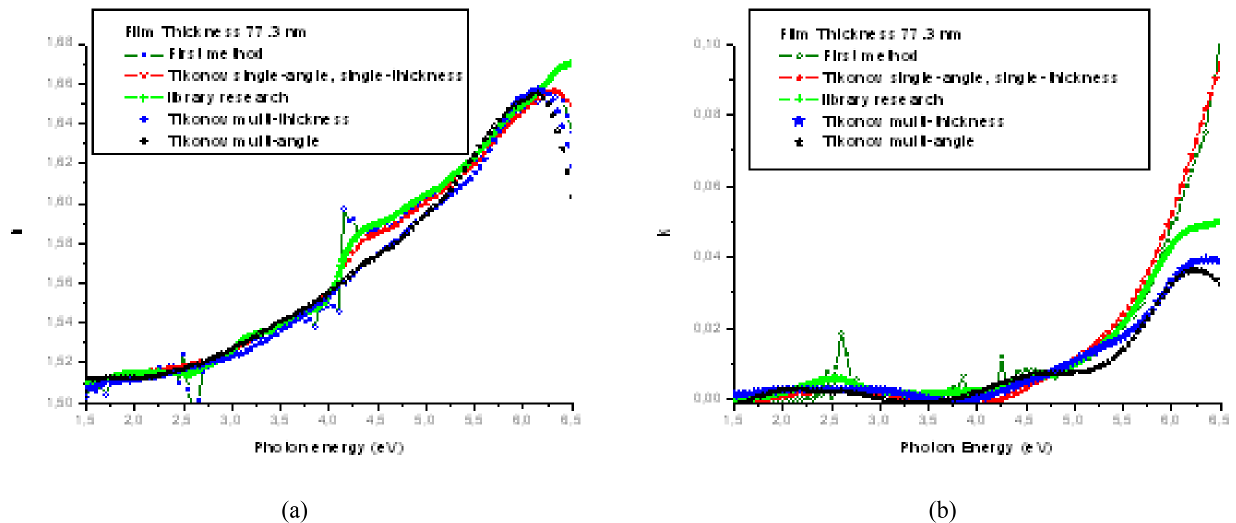


Figure 2: (a) *n* vs. photon energy and (b) *k* vs. photon energy (b) for the 77.3 nm thick material

The first method (local optimization method, LO) gives noisy curves, particularly at the boundary of the energy domain. This is easily explained as no regularization is applied to the simulation results and hence any noise in the measurements will give noise in $n$ and $k$. For the 77.3 nm thickness, the jumps in the curve are numerous (see Figure 2). On the other hand, few jumps in the curve are reported for the 103.1nm thick film. This shows that the presence of local minima is specific to the combination of the film thickness and the incident angle. For instance at 77.3nm with an incident angle of 70°, the deepest local minimum for $n$ is located at 4.2 eV. In order to try to improve the accuracy of the process, we used Tikhonov regularization at single angle and single thickness. Results show good correspondence with the no regularization case but this time $n$ and $k$ are smoother. Nevertheless local minima are not avoided but smoothed out (see Figure 2). For the Tikhonov multi-angle optimization we use three different angles, thus we have $6N$ measurement data. For the Tikhonov multi-thickness optimization we have used $4N$ data, such as $Is^m$ and $Ic^m$ for both 103.1 nm and 77.3 nm. Those two curves (multi-angle and multi-thickness) are quite close to each other. No local minima on n and k are detected compared to the single angle regularization (Tikhonov or Library search). The over determination strategy of using additional input appears to be more robust (see Figures 1 and 2). The library search is quite close to the multi-angle and multi-thickness Tikhonov regularization but does not always give smooth parameter variation (see Figure 2). For the 103.1 nm thick film, the library search is quite close to multi-angle and multi-thickness regularization. This is not exactly the case for 77.3 nm where the regularization does not fully prevent from being attracted by the local minima at 4.2eV. Tikhonov regularization for single-angle and single-thickness behaves in a very similar way than the library search but the main difference is that the final results of the Tikhonov regularization will be dependent of the initial conditions and this issue has to be handled with great care. One advantage and not the least of the library research is that no initial conditions are needed. But it is clearly not as robust as multi-angle and multi-thickness regularization (see Figure 2b). To conclude, the best values of the optical index of thin resist films are obtained with the multi-angle or multi-thickness Tikhonov regularization.

## 5. CONCLUSION

The problem we have studied in this paper is well known in the literature of non-linear inverse problems. Different techniques have been proposed and tested by different authors. Due to the nature of ill-posed problems, many difficulties arise during the optimization stage. Moreover, the problem we tackle with concerns thin film characterization, which means that the measurements are nanometric and can be very noisy. We thus propose different methods to solve this issue, where the unknowns are the thickness, the refractive index and the coefficient of extinction of thin films deposited on a Silicon substrate.

The first method we use is a classical local optimization routine which aims at minimizing the errors between measured data and theoretically computed values. Due to the lack of a regularization component, the parameter values obtained can be noisy with this technique. As a second method, we use in addition to the first one a smoothing technique: Tikhonov regularization. Cross validation is used to choose the best value of the regularization coefficient, which role is to smooth out the parameter curves. To be more accurate, we also increased the number of independent measurements using multi-angle and multi-thickness measurements. We tested this strategy using also Tikhonov regularization. As a third method the library search is used, which consists in choosing the parameters which fit the best to the measured data. The search is made in a dense library of theoretically pre-computed $Is$ and $Ic$ curves.

We tested all these methods for different materials of different thicknesses. We found out that the multi-angle and multi-thickness Tikhonov regularization seem more realistic, more robust, include less noise and both are very close to each other for the same material. The first method of local optimization can give noisy curves due to the lack of a regularization component. The library search can be quite close to the multi-angle and multi-thickness Tikhonov regularization for some tests but some unsmoothed parts remain in the curves. This method is still under development. One of its major advantages is to provide a "visual" access to the cost function therefore being able to spot the potential local minima. Increasing the number of input data is a very efficient way to improve the robustness of the optimization. We can conclude that the use of multi-angle and multi-thickness Tikhonov regularization results in more accurate and robust results. However for the large data sets, the cross-validation used to find the best regularization level for the Tikhonov approach can be very time consuming. In this case the library search seems to be quite efficient for its speed and its quality of resulting curves. However, like in every optimization process, the "correct" result is not known and only a careful analysis of the results and robustness tests can provide a better confidence in the output results.

# REFERENCES

[1] Tompkins, H. and Irene, E. (Ed.), [Handbook of ellipsometry], William Andrew Publisher, (2005).
[2] Tarantola, A., [Inverse problem theory and methods for model parameter estimation], SIAM Publisher, (2005).
[3] Haber, E., "Numerical strategies for the solution of inverse problems", Ph.D. dissertation The University of British Columbia, (1997).
[4] Hansen, C., [Rank-Deficient and Discrete Ill-Posed Problems], SIAM Publisher, (1998).
[5] Busby, H. and Trujillo, D., "Optimal regularization of an inverse dynamics problem", Computers & Structures 63(2), 243-248 (1997).
[6] Krawczyk-Stando, D. and Rudnicki, M., "Regularization parameter selection in discrete ill-posed problems -- The use of the U-Curve", International Journal of Applied Mathematics and Computer Science 17(2), 157-164 (2007).
[7] Bobro, V., Mardezhov, A. and Semenenko, A., "On the solution of incorrect inverse ellipsometric problem", Proc. SPIE 3485, 354-358 (1998).
[8] Rosa, R., "The inverse problem of ellipsometry: a bootstrap approach", Inverse Problems 4, 887-900 (1988).
[9] Tonova, D. and Konova, A., "Characterization of inhomogeneous dielectric coatings with arbitrary refractive index profiles by multiple angle of incidence ellipsometry", Thin Solid Films 397, 17-23 (2001).
[10] Kone, I., "Caractérisation optique de matériaux et métrologie optique d'objets nanométriques par la résolution d'un problème inverse", Master Dissertation Université Joseph Fourier, Grenoble, (2006).
[11] Coleman, T. F. and Li, Y., "On the convergence of interior-reflective Newton methods for nonlinear minimization subject to bounds", Mathematical Programming 67(1), 189-224 (1994).
[12] Coleman, T. F. and Li, Y., "An Interior Trust Region Approach for Nonlinear Minimization Subject to Bounds", SIAM Journal on Optimization 6(2), 418-445 (1996).
[13] Tikhonov, A. and Arsenine, V., [Solution of Ill-posed Problems], Winston and Sons Publisher, (1977).
[14] Soulan, S., Besacier, M., Leveder, T. and Schiavone, P., "Real-time profile shape reconstruction using dynamic scatterometry", Proc. SPIE 6518(1), 65180W (2007).
[15] Soulan, S., "Développement des techniques de scatterométrie pour le suivi de procédés de gravure plasma", Ph.D. Dissertation Université Joseph Fourier, Grenoble, (2008).