Journal of Global Optimization **19:** 383–401, 2001.

A Combined Topographical Search Strategy with Ellipsometric Application

O. POLGÁR, M. FRIED, T. LOHNER and I. BÁRSONY
Research Institute for Technical Physics and Material Science, H-1525 Budapest, Hungary (e-mail: polgaro@mfa.kfki.hu)

(Received 10 September 1999; accepted in revised form 14 December 2000)

Abstract. A comparison deals with the advantages and disadvantages of the classical random-base, exhaustive and gradient searches and presents a precise local search combined global search control strategy including a new, systematic point selection which makes possible the escape from local minima by time. As a demonstration electrochemically etched porous silicon (PS) samples were investigated by spectroscopic ellipsometry (SE). The evaluation process (a global optimisation task) was made in different ways to see the difficulties and the differences among the evaluating possibilities. The new, topographical search (named Gradient Cube search) was compared with some classical methods (Grid search, Random or Monte-Carlo search, and Levenberg-Marquardt gradient search) and with two more complex algorithms (Genetic Algorithms and Simulated Annealing) by evaluating real measurements. The application results prove that the classical methods have difficulties to give enough reliability and precision at the same time in global optimisation tasks if the error surface is hilly. There is therefore a hard need of escaping from local minima, and a need of a systematic evaluation to avoid the uncertainty of random-base evaluation. The Gradient Cube search is an effective, systematic hill-climbing search with high precision and so it can be useful in ellipsometry.

Key words: Topographical search, escaping from local minima, global optimisation, minimum searches, Ellipsometry

1. Introduction

The technique of ellipsometry [1] is recently in a renaissance driven by an ever-increasing demand for rapid, non-destructive analysis of surfaces and thin films. Ellipsometry enables the determination of optical constants of materials with high accuracy and can therefore help to solve a wide variety of problems in different disciplines. The principle of ellipsometry was established about a hundred years ago (Drude, 1889, 1890), but the technique has been efficient by used only recently in parallel with the development of rotating element ellipsometers and the spreading of microcomputers. The reason is simple; in lack of inverse equations there is no possibility of directly finding the proper parameters of the optical model of the sample from the measured spectra. A complicated multi-layer structure is modelled as a system built up of plan-parallel thin films (see Figure 1), that can each consist of a mixture of two (or more) different materials. The difficulties arise

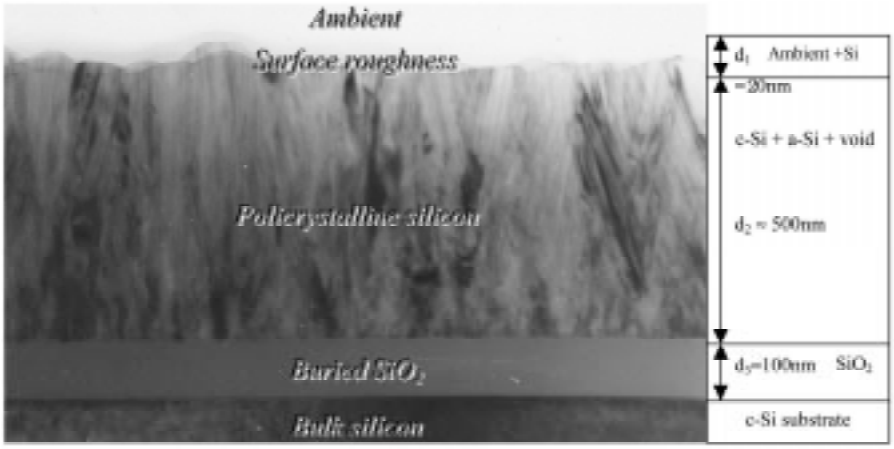


Figure 1. A multi-layer structure and its optical model. The optical model doesn't contain the effect of the surface roughness.

from the physical and mathematical background, as non-linear, iterative methods are required to find the proper parameters of the optical model, because there is no inverse equation to determine them directly from the spectra.

The theory of polarimetric reflectometry (i.e. ellipsometry [1]) is based on the change of the polarisation state of the light reflected on the sample (see Figure 2). The change of the amplitude can be measured as φ (or tan (φ)), and the change of the phase as Δ (or cos (Δ)). Since the polarisation changes depend on the wavelength (energy) of the incident light, in spectroscopy (or in spectroscopic ellipsometry (SE)) about 100, different wavelengths are used in the ultraviolet (UV), visible (VIS) and near infrared (NIR) range. Hence the measured data will consist of two spectra, tan (φ) and cos (Δ) versus the wavelength, describing the reflection properties of the surface. These, in themselves give no direct information about the multilayer structure or even the composition of the sample. To obtain this piece of information an optical model of the surface must be assumed [2] (see Figure 1). Usually an idealised optical model of the measured sample can be created from the history of the sample (a-priori information). Most optical models use flat semi-infinite substrates with one or more laminar adherent layers of uniform thickness on the surface. All interfaces are assumed to be sharp, and all layers are assumed to be composed of optically isotropic materials by using the Bruggemann effective medium theory, [1]:

$$f_a \frac{\bar{\epsilon}_a - \bar{\epsilon}}{\bar{\epsilon}_a + 2\bar{\epsilon}} + f_b \frac{\bar{\epsilon}_b - \bar{\epsilon}}{\bar{\epsilon}_b + 2\bar{\epsilon}} = 0 \tag{1}$$

where f_a and f_b are the concentrations of components a and b materials $\bar{\epsilon}_a$, $\bar{\epsilon}_b$, $\bar{\epsilon}$ are the dielectric functions of a, b and the mixture.

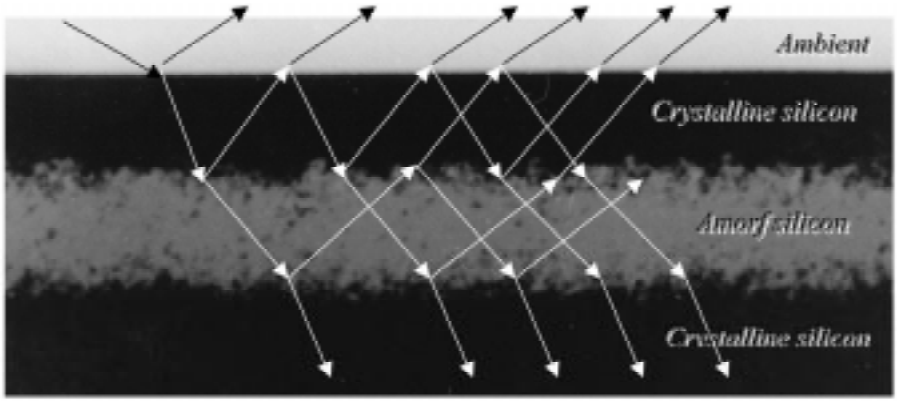


Figure 2. The light reflection on a multi-layer sample. The ways of the light are only demonstrative.

The goodness of the optical model is estimated by the unbiased estimator:

$$\mathrm{RMSE} = \frac{1}{2N - P - 1}\sqrt{\sum_{j=1}^{N}\{(\tan\Psi_j^s)^2 + (\cos\Delta_j^m - \cos\Delta_j^s)^2\}} \tag{2}$$

where RMSE is the Root of Mean Squared Error,

N is the number of the points of the spectra,
m is the measured and s is the simulated (calculated from the optical model) spectra,
and P is the number of the parameters in the optical model

Finding the minimal value of the RMSE, changing the variable parameters of the optical model, leads to a classical, global optimising (error minimising) task. The classical solutions offer either high accuracy with strong limitations (gradient descent methods [4, 10] work properly only if the starting point is in the monotone decreasing neighbourhood of the global minimum, see Figures 3 and 4), or hill climbing ability with lower accuracy (random-base: Random or Monte-Carlo search, Simulated Annealing (SA) [9], Genetic Algorithms (GA) [6, 7] etc.) or more reliability but high time consumption (Exhaustive or Grid search). Combining these searches in a sequence of either a random or a grid pre-search and then a gradient method may give better properties, but the strategy of their alternate use is not well defined, because the critical RMSE sophistically depends on the optical models and the measurements. The task of the pre-searches is to find the monotone decreasing neighbourhood of the global minimum, from where a gradient descent method surely finds the solution. Either using more complicated optical model by ascending the number of the parameters (higher number of search space dimension) or increasing either the resolution (reducing the grid distances) or the size of the search space, combinatorial explosion takes effect. Local minima and the problem of deciding whether a minimum is local or global in the lack of knowing the RMSE

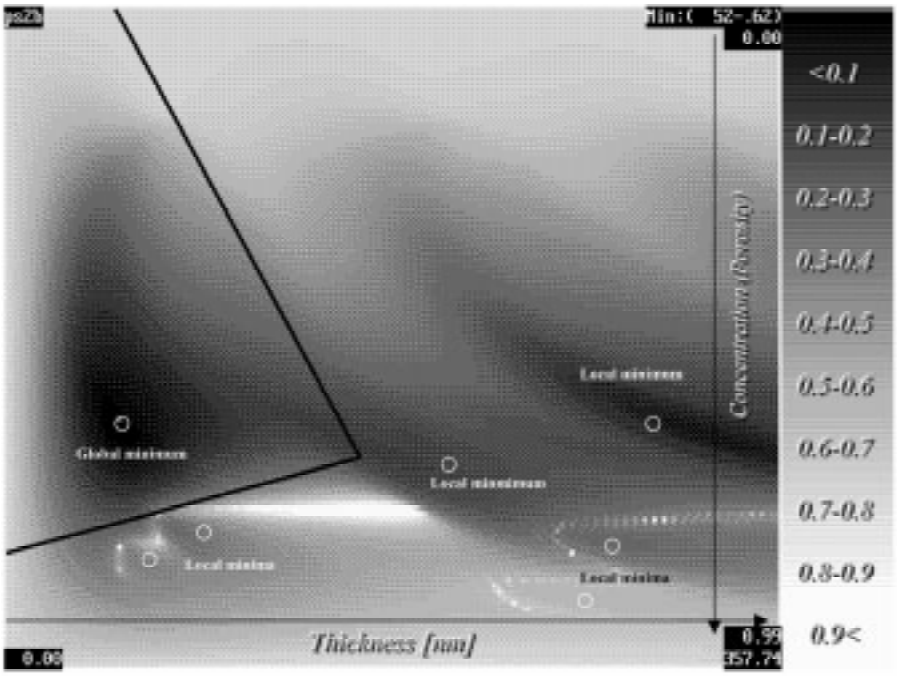


Figure 3. Two-dimensional error 'surface' of porous silicon layer. The global minimum and several local minima can be seen in a view of a porous silicon layer (see below). The X axis means the thickness of the ps. layer and the Y axis means the porosity; i.e. the void concentration in the crystalline silicon. The darker grey points mean lower RMSE. The small white circles mean the global and the local minima. If the gradient descent method started somewhere beyond the black lines it would slip in a wrong valley.

Table 1. Optical model

PS366	1. Component	2. Component	Concentration	Thickness
1. film	SiO_2	–	–	5
2. film	Crystalline Si	Void	0.75–0.25*	100
Substrate	Crystalline Si	–	–	∞

*The 2. film was divided into 10 sublayers. A linear decreasing concentration profile was used. It was described in two variables, i.e. the top and the bottom concentrations.

of the solution, cause also difficulties. The random feature causes uncertainty because different evaluations, even with the same random-base method, may give different results. A systematic search gives more reliability. Besides, the problem of how long (how many cycles) the pre-search should be run in the lack of knowing the range of the critical RMSE also causes difficulties and decreases the reliability of the evaluation.

Table 2. Random pre-search and after Gradient method compared to Gradient Cube search

PS366				Grid pre-search + gradient search			Random pre-search + gradient search		Gradient cube search	
4 parameter-optical model	Starting point	Low limit	High limit	Step	Grid search	Gradient search	Random search	Gradient search	Initial edge	Gradient cube search
Cycles	0				3800	3	2000	7	–	6 cube + 32 gradient steps
Comp. Time	0				4.75h	+105s	2.5h	+246s	–	2h
RMSE	0.326	–	–		0.064	0.060	0.108	0.086	–	0.042
1. film thick. [nm]	20	0	20	5	0		0.84	2.48	0.33	3.86
2. film thick. [nm]	100	50	500	25	125		107.16	101.16	3.33	269.22
2. f. upper conc.	0.75	0.25	1	.1	0.95		0.49	0.48	0.01	0.47
2. f. lower conc.	0.25	0.0	0.5	.1	0.3		0.24	0.05	0.01	0.30

PS366	Gradient cube search	Grid search	Random search
RMSE after pre-search	0.042	0.064	0.108
RMSE after gradient search		0.060	0.086

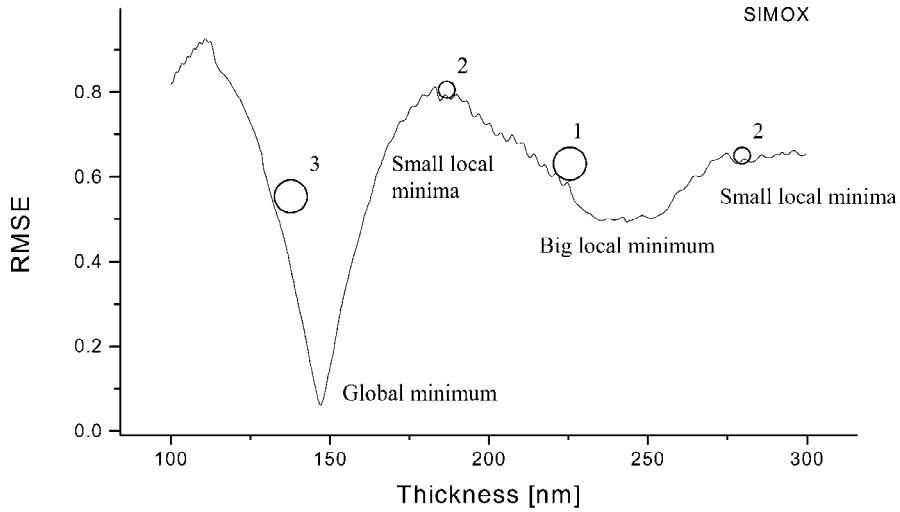


Figure 4. One-dimensional error surface of a three-layer SIMOX structure [12]. The local minima can be bigger and smaller depending on the so-called momentum of the search, which can be modelled with the greatness of the balls. The smaller balls (2) have smaller momentum and may slip in smaller minima, meanwhile the bigger ball (1) is able to roll over. In some kind of gradient methods the idea of the momentum are used.

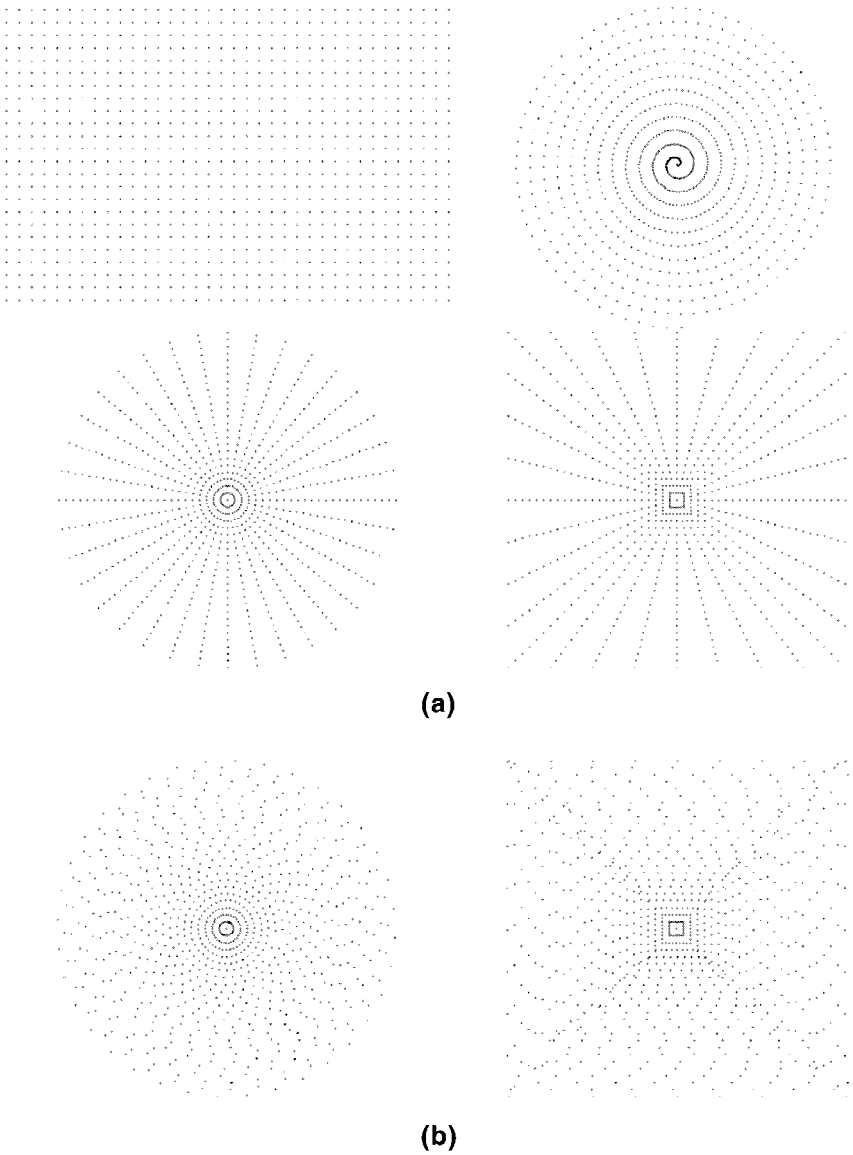
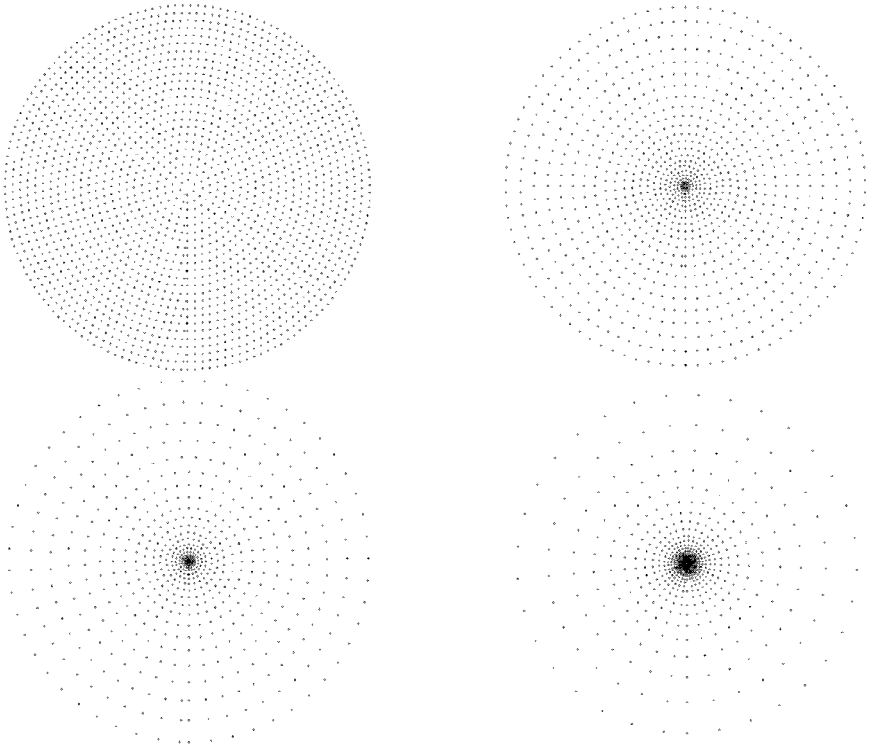


Figure 5. Concentric two-dimensional strategy shapes. (a) 2 dimensional grid, spirals, spheres, and cubes with constant parameter steps (linear ascending radius); (b) dimensional randomly rotated spheres and cubes



(c)

Figure 5. (c) homogeneously selected circle-points applying first, second, and third order ($|r_i|$ = i^{n}*r_0) and exponential (r_i | = r$_0$∗ e$^{i\tau}$) radius function where τ=5

Table 3. Optical model

PS372	1. Component	2. Component	Concentration	Thickness
1. film	Fine-grain poly-Si	Void	0.85	20
2. film	Fine-grain poly-Si	Void	0.73	16
3. film	Fine-grain poly-Si	Void	0.61	70
Substrate	Crystalline Si	–	–	∞

Because the above mentioned problems there is a hard need of an accurate, systematic optimisation algorithm in ellipsometry, which is able to escape from local minima providing higher reliability. The unique kind of searches with satisfied precision are the gradient-base methods, but possible local minima reduce the reliability. The solution may be a hybrid control strategy, combining the high precision gradient-base local optimisation with a systematic global search. The control strategy should provide a hill climbing ability in order to give higher reliability

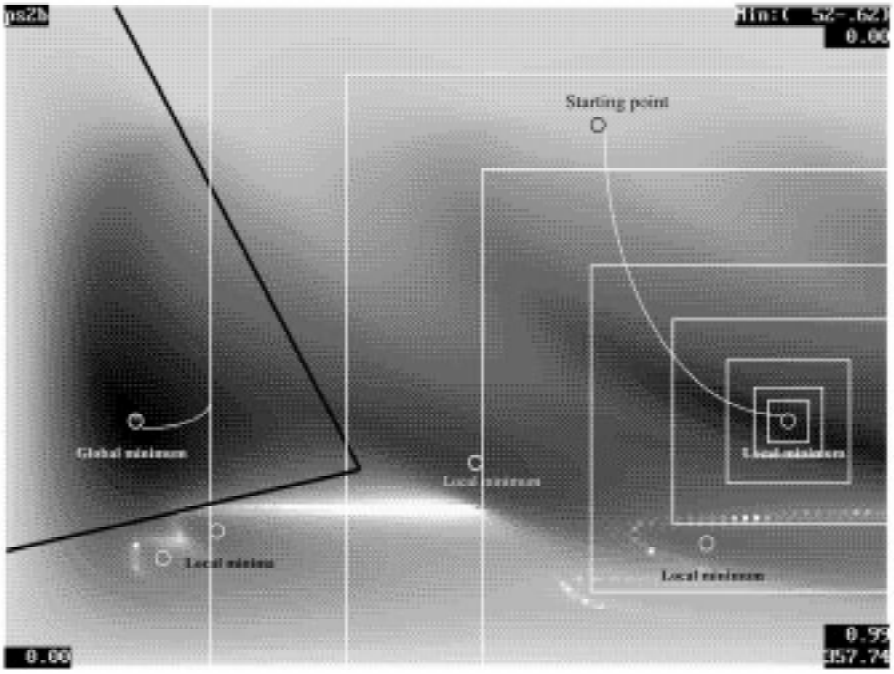


Figure 6. Control strategy. From the starting point (black circle) the gradient search will find a deep local minimum. In the next steps the concentric cubes will be expanded by increasing the edge, and in parallel the costs of the centre are increased, until a point with better criteria function is found. From this point the gradient search will be started again and find the global minimum. The figure is only a demonstration, not a real example. For better understanding the area of the cubes and the curve of the gradient movement can be seen instead of the examined points.

Table 4. Random pre-search and after Gradient method compared to Gradient Cube search

PS366				Grid pre-search + gradient search			Random pre-search + gradient search		Gradient cube search	
7 parameter optical model	Starting point	Low limit	High limit	Step	Grid search	Gradient search	Random search	Gradient search	Initial edge	Gradient cube search
Cycles					86400	2	130000	5	–	8 cube + 55 gradient steps
Comp. Time					30h	43s	44h	108s	–	3.7h
RMSE	0.178				0.096	0.074	0.053	0.041	–	0.029
1. film thick. [nm]	5	0	20	5	15	14.8	2.34	2.56	0.33	3.45
2. film conc.	0.85	0.65	0.95	0.1	0.75	0.72	0.83	0.82	0.01	0.821
2. film thick. [nm]	20	10	60	10	20	19	31.49	29.53	1	35.308
3. film conc.	0.73	0.5	0.9	0.1	0.8	0.78	0.62	0.64	0.01	0.615
3. film thick. [nm]	16	10	60	10	40	40	22.14	28.28	1	30.102
4. film conc.	0.61	0.4	0.7	0.1	0.6	0.63	0.50	0.50	0.01	0.481
4. film thick. [nm]	70	10	60	10	30	29.9	50.97	50.62	1	45.690

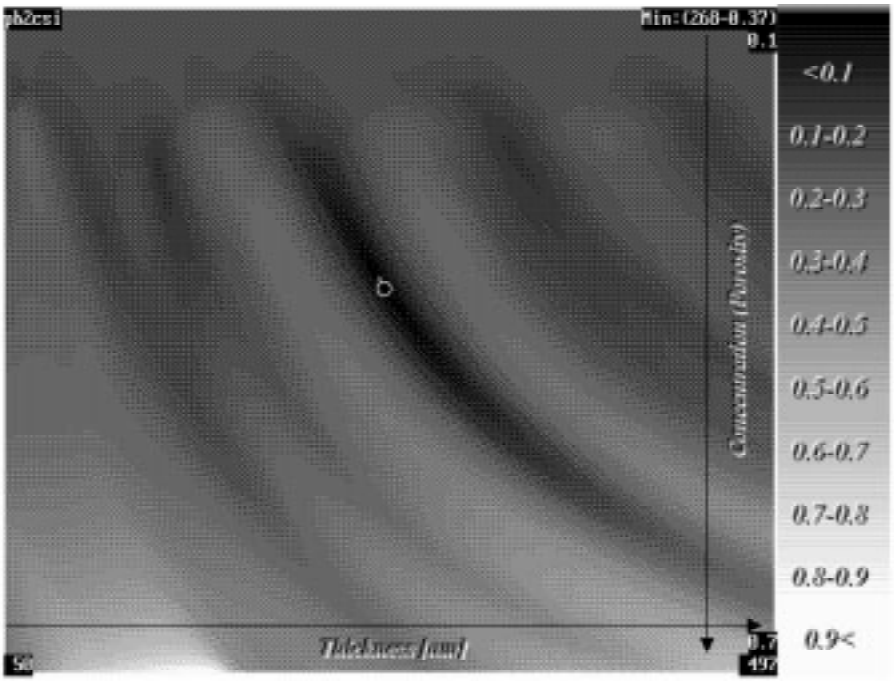


Figure 7. A two-dimensional error surface of PS366 search space. The global minimum (white circle) and several local minima (somewhere in the darker valleys) can also be seen in the figure.

Table 5. Optical model

1. Component	2. Component	Starting thickness [nm]	Best thickness found [nm]	Thickness limits [nm]	Starting porosity	Best porosity	Porosity limits
SiO_2	–	5	–	–	–	–	–
Fine grain poly-Si	Void	100	48.1	0..200	0.5	0.68	0..1
Fine grain poly-Si	Void	100	40.8	0..200	0.5	0.44	0..1
Fine grain poly-Si	Void	100	41.6	0..200	0.5	0.26	0..1
Fine grain poly-Si	Void	100	95.4	0..200	0.5	0.20	0..1
Fine grain		100	9.7	0..200	0.5	0.59	0..1
Crystalline Silicon	–	∞	∞	–	–	–	–

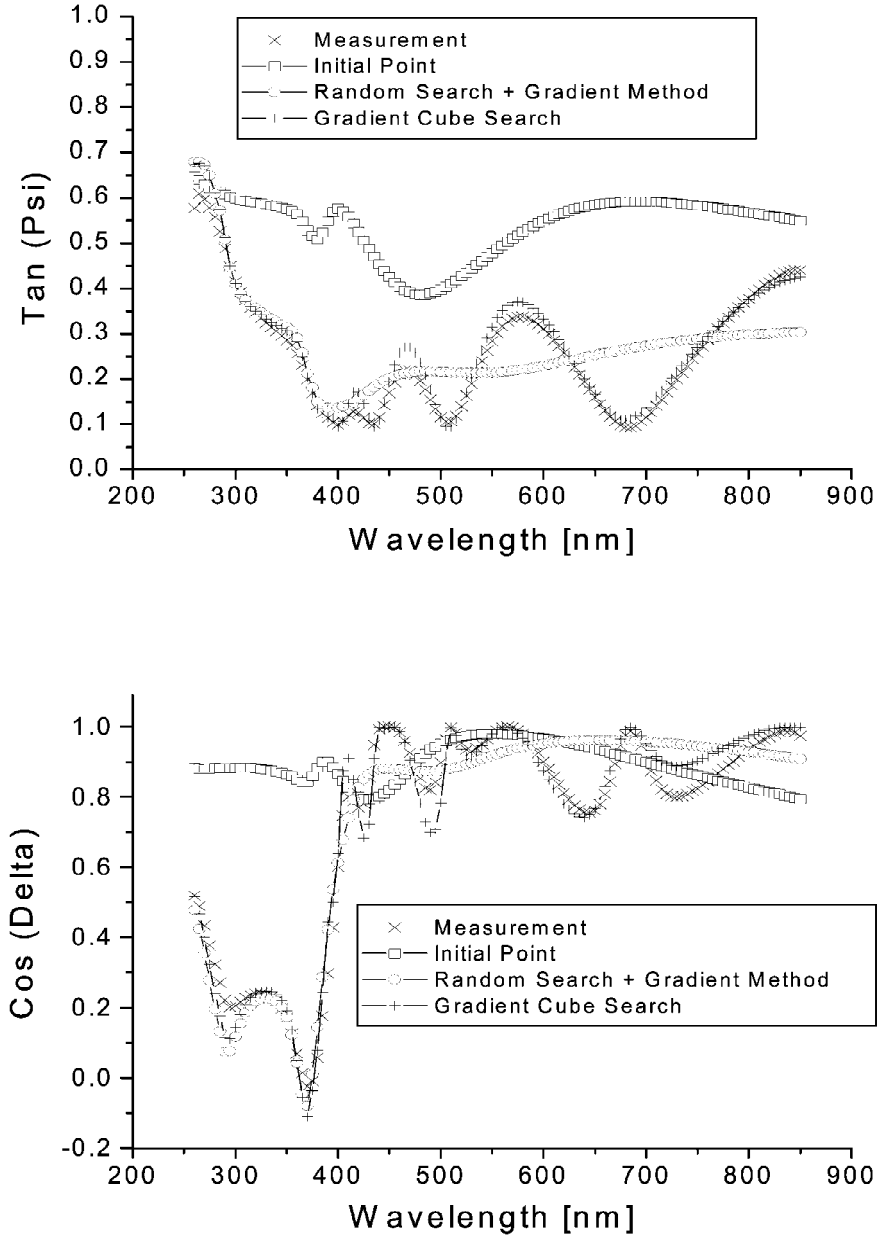


Figure 8. Spectra of PS366.

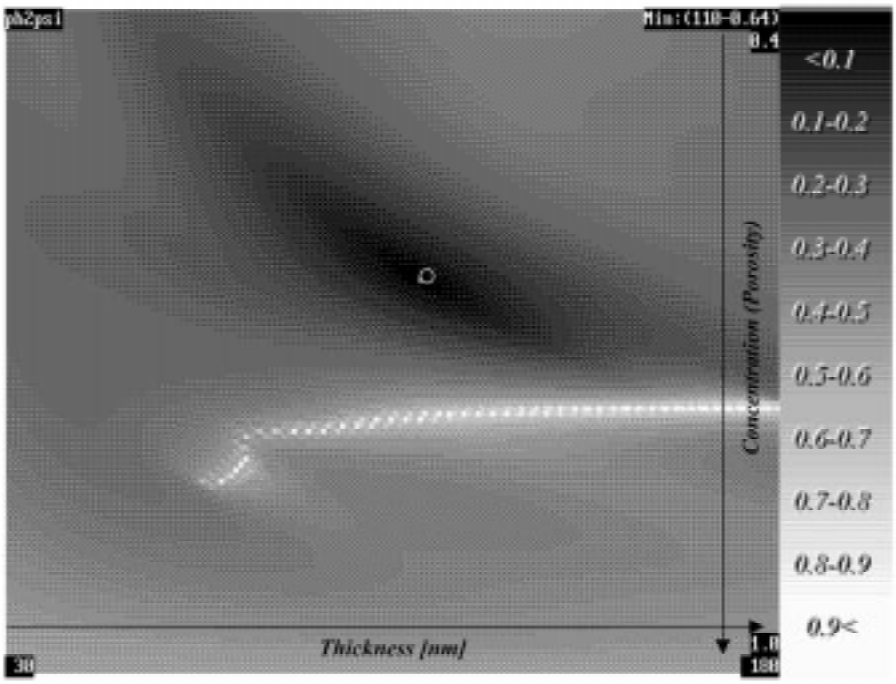


Figure 9. A two-dimensional error surface of PS372 search space. The global minimum and a local minimum (lower right corner) can also be seen in the figure.

Table 6. Random pre-search and after Gradient method compared to Gradient Cube search

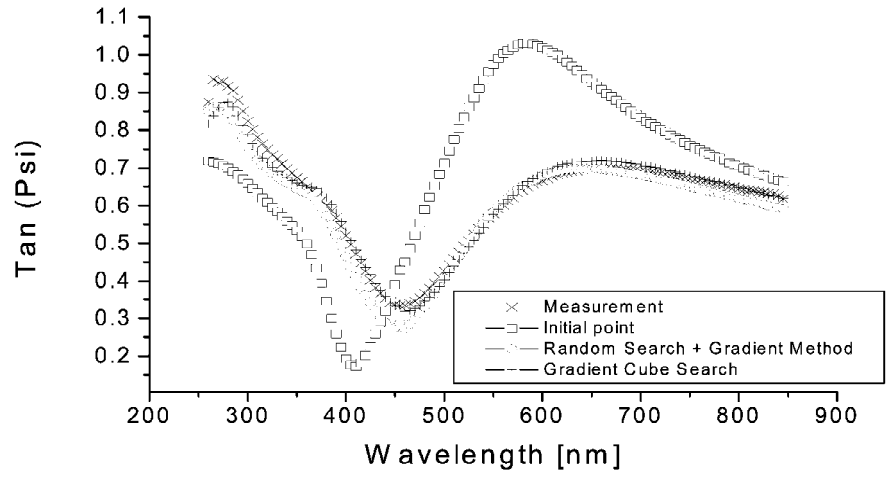
PS990	Gradient cube search	GA	SA	Random search
RMSE after 12 hour-pre-search	0.041	0.056	0.078	0.095
RMSE after gradient search		0.052	0.065	0.086

even if the computation times increase. A hybrid control strategy (named Gradient Cube search), including a new, systematic point selection, was tested to evaluate ellipsometric measurements.

2. Gradient Cube Search

USING EXTENDED CRITERIA FUNCTIONS

The hybrid control strategy (detailed later) uses two different, local and global, point selections alternatively. The local one, a gradient descent method (Levenberg-Marquardt, modified by Fletcher [4, 10]) finds always the local minimum of the



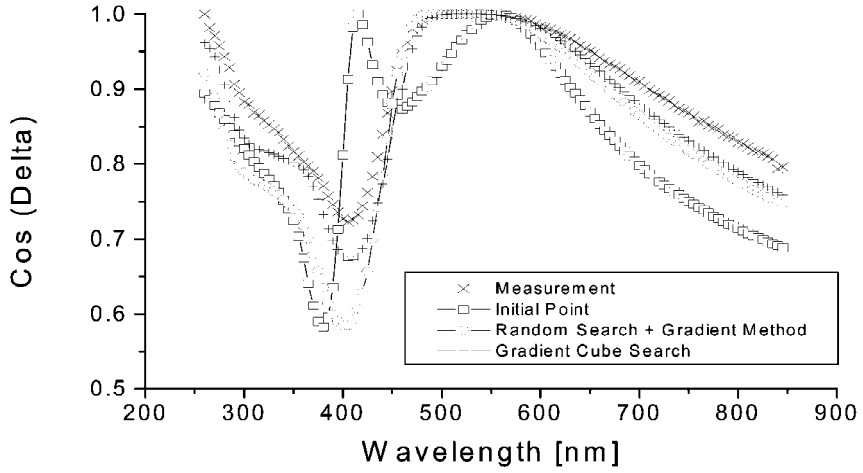


Figure 10. Spectra of PS372.

neighbourhood with high precision. From the local minima an expanding systematic point selection (detailed later) guaranties the escaping ability. The control strategy always selects the local minimum (and its neighbourhood) with the best criteria function (see Equation (3)), which contains the RMSE of the centre and an increasing cost component representing its neighbourhood.

$$f_1(\underline{x_c}) = C_h * h(\underline{x_c}) + C_c * c(\underline{x_c}) = C_h * \text{RMSE}\,(\underline{x_c}) + C_t * \sum_{i=1}^{n_x} t(\underline{x_i}) + C_e * \sum_{i=1}^{n_x} \text{RMSE}\,(\underline{x_i}) \qquad (3)$$

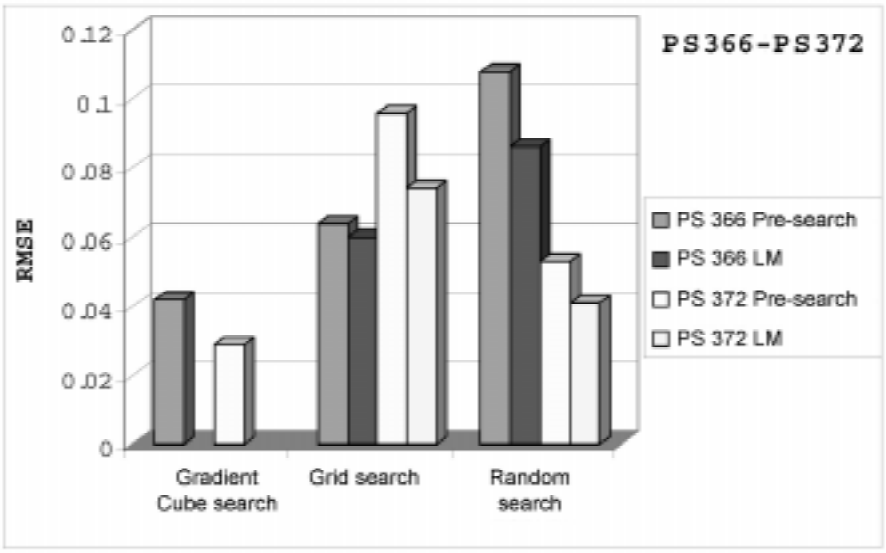


Figure 11. Results of PS366 and PS372.

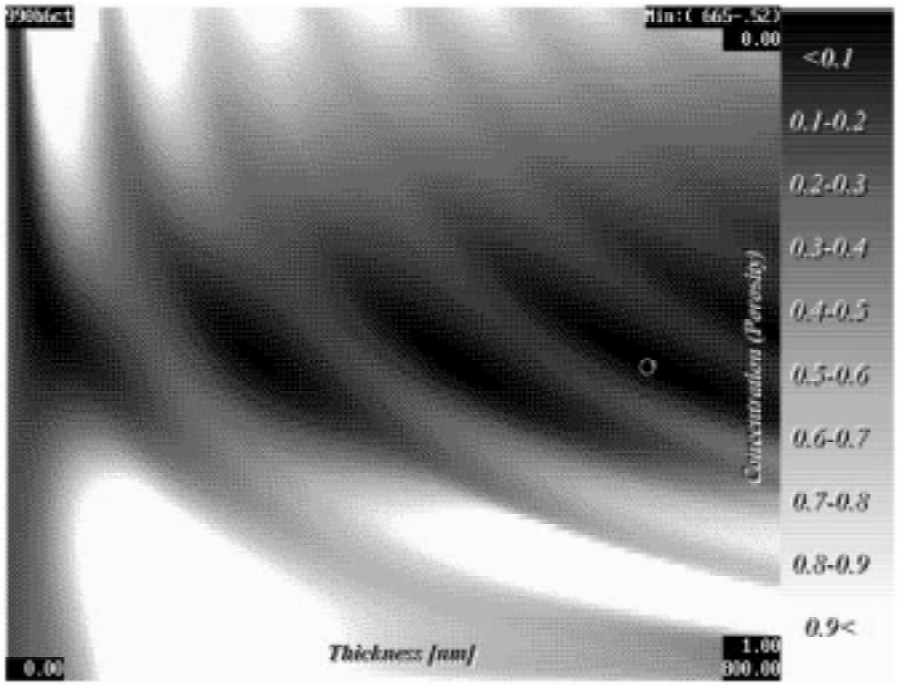


Figure 12. A two-dimensional error surface of PS990 search space. The global minimum and several local minima can be seen in the figure.

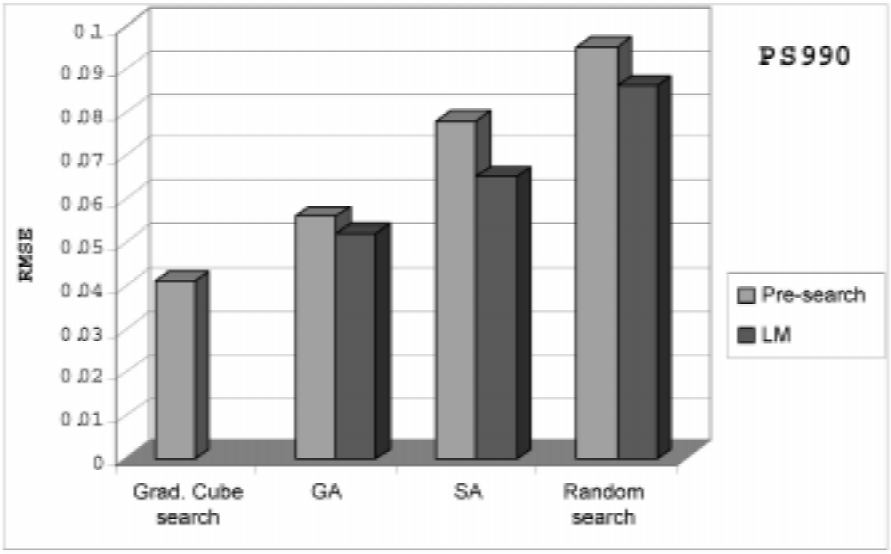


Figure 13. Results of PS990. The starting points of the LM gradient searches determined by different pre-searches (GA, SA, Random search). The computation time of the pre-searches was 12 hours. The Gradient Cube search includes the gradient search therefore it gives only one result.

where $f()$ means the criteria function,
$h()$ means the expected distance from the solution, supposing the RMSE of the solution is zero,
$c()$ means the accumulated costs of the centre,
x_c is the centre of the cube, x_i are the n_x examined points belonging the same centre, C_h, C_c, C_t and C_c are experimental weights. C_t was zero in the tests because the computation times of points are equal in ellipsometry.

The $h()$ function estimates the goodness of the centre. The extended component, named cost function $c()$ combines two tasks. It partly estimates the goodness of the neighbourhood, recording to the already examined points belonging to the area, and partly gives preference to less known areas. So the extended criteria function always prefers the most hopeful area. The areas are not systematically selected before starting the search but during the systematic point selection. Each examined point belonging to an extending cube can be a new centre of a new area if its RMSE is better than the worse centre stored in the memory. The criteria function of a new point is only its RMSE first, because the cost component is zero (no examined point belonging to the new centre). If all points are memorised another type of criteria function can be used (see Equation (4):

$$f_2(\underline{x_c}) = C_h {}^{*}h(\underline{x_c}) + C_c {}^{*}c(\underline{x_c}) = C_h {}^{*}\mathrm{RMSE}\ (\underline{x_c}) + C_t {}^{*}\sum_{i=1}^{n} e^{\frac{-|x_i - \underline{x}|}{r^{*}d_{\max}}} * \mathrm{RMSE}(\underline{x_c}) \tag{4}$$

Here each examined point belongs to each centre reverse proportionally with the distance between the point and the centre. In Equation (4) an exponential proportionality can be seen. In the comparison tests the first function (Equation 3) was used.

POINT SELECTION OF GLOBAL SEARCH

The expanding point selection strategy systematically selects the points lying closer to the centre at first, then the points with higher and higher distances from the centre. The classical grid search and three new point selection strategies can be seen in Figure 5a. Because of the easiest coding the cube-search was applied in the tests. Avoiding the 'avenues' the areas of spheres and cubes can be rotated as can be seen in Figure 5b. If the points are selected in equal distances from each other, the steps can be determined as the average distances of neighbouring two areas (circles, or cubes). See Equation (5) and Figure 5b.

$$S_i = 1/2^*(r_{i+1} - r_i) + (r_i - r_{i-1}) = 1/2^*(r_{i+1} - r_{i-1}) \tag{5}$$

where r_i are the radius of the i^{th} circles

$$\text{The number of points in a circle :} \quad N_i = 2^* r_i^* \Pi / S_i = 4^* r_i^* \Pi / (r_{i+1} - r_{i-1}) \tag{6}$$

$$\text{Hence the unit degree :} \quad 360/N_i = 360/(4^*\Pi)^*(r_{i+1}/r_i - r_{i-1}/r_i) \tag{7}$$

Replacing the linear ascending radius (in cubes: edge) function with higher order functions, close to the centre smaller and far away from the centre larger steps can be used in order to give bigger chance to the points being closer to the centre. The better points are closer to the centre with bigger probability in practice. The optimal radius-function may depend on the measurement and the applied optical model, which determine the search space. Combinatorial and exponential functions were examined. See Figure 5c.

CONTROL STRATEGY

The alternative use of local and global searches (see Figure 6):

- *Gradient movement (local search)*

From the starting point (with estimated parameters) a gradient method finds first the closest local minimum, and the local minimum will be stored in the memory with its zero costs criteria function. Then the combined search strategy continues with the hill climbing point selection.

- *Hill-climbing point selection (escaping from local areas)*

After a local minimum found, the combined search continues with the systematic hill climbing point selection, i.e. expanding concentric cubes. The centre of the cubes is always a local minimum with the best (least) criteria function. Examining the area of the cubes the costs of the centre point is accumulating, and so the criteria function is increasing (becoming worse). The points are stored in the memory sorted by their criteria functions. A duplicate point examination, caused by the crossing areas of different cubes, is not possible because of the use of a point register look-up-table. Each newly examined point has zero costs until being a centre. If a point with best criteria function is found:

a) If the point is new (not stored in the memory), a gradient search will take place again to move the point into the local minimum of its neighbourhood (gradient movement), and instead of storing the found point, the local minimum will be stored only if it is also new. The edge of the next cube will be the initial one (the smallest) again.

b) If a previously stored local minimum becomes the best (because the costs of the others become higher) its extension continues with a bigger edge.

Pseudocode

Enter the initial point and step of each parameter
Enter the limitation of the search (time, points, $RMSE_{crit}$ etc.)
Start gradient search
Memorise the local minimum with zero costs
REPEAT
  Select the best point (bp) and its attributes (edge of the actual cube, position of the actual point on the area, costs, f_{crit}) from the memory regarding the criteria functions
  Select the next point (bp_i) on the area of the actual cube of the point bp if it isn't yet examined
  IF the last point of the area examined THEN
    Increase the edge of the cube and continue with the first position again
  END IF
  Compute the RMSE of the point
  Increase the costs and so the criteria function (f_{crit}) of bp
  IF $f_{crit}(bp_i) < f_{crit}$(worst) THEN
    Start a gradient search
    IF the found local minimum (pl) isn't yet stored in the memory THEN
      Put it in with zero costs, initial edge and first position (exchange with the last one if the memory is full)
      Make the memorised points in order by their criteria functions
    END IF
  END IF
  UNTIL limitation (time, points, $RMSE_{crit}$ etc.)

3. Results

In order to demonstrate the method, described before, 3 different measurements comprising 3 optical models were used. Each sample was electrochemically prepared porous silicon (PS) with different porosities [5]. Porous silicon can be produced from crystalline silicon by electrochemical dissolution in hydrofluoric acid. PS was formed on p-type silicon wafers by anodisation in ethanol containing aqueous electrolyte with HF contents. The duration of the etching is typically several seconds or minutes. The ellipsometric measurements were made by the use of a rotating analyzer ellipsometer. Two optically different types of silicon forms, a bulk-type silicon (c-Si) and a fine-grain polycrystalline silicon with enhanced absorption in the grain boundaries (p-Si) needed to be mixed with voids in the appropriate ratio, and the PS had to be divided in depth in several different sections in order to obtain the best fit. The first PS measurement (named PS366) was modelled first as a mixture of void and crystalline silicon. See the optical model in Table 1 and a two-dimensional error surface of the search space in Figure 7. The spectra and the results can be seen in Figure 8 and 11.The complex refractive index of the layer was calculated by Bruggeman effective medium approximation [1] by using void and crystalline silicon as end-points, and a linear decreasing concentration profile was applied. Evaluating the second measurement (named PS372) the PS was considered as a three-layer structure of mixtures of fine-grain polycrystalline silicon [8] and void. (See the optical model in Table 3 and a two-dimensional error surface of the search space in Figure 9.) The spectra and the results can be seen in Figures 10 and 11. The third measurement (named PS990) was modelled as a multi-layer structure build from thin, parallel films of mixtures of fine-grain polycrystalline silicon and void in order to consider the change of concentration in depth. See the optical model in Table 5 and a two-dimensional error surface of the search space in Figure 12. The results can be seen in Figure 13. In this case the SiO_2 film on the top of the sample was also considered.

The search spaces were the same when the Grid search, the Random search and GA were used. SA and the Gradient Cube search can move the centre points anywhere, so that they work with unlimited search spaces. Avoiding the combinatorial explosion, the grid distances had to be selected much larger in the Grid search than the initial edges of cubes in the Gradient Cube search, which results bigger errors.

Started at the initial points the Levenberg-Marquardt gradient descent method stopped on high level local minima in all cases, which proved that the presence of local minima and the need of more complex search algorithms. Started at low-error points found by pre-searches the L-M gradient method slipped in local minima in all cases. However the Gradient Cube search required approx. 2 hours in the first case, and less than 4 h in the second case and found better points (with lower RMSE) which can be seen in Figure 8. The simulated curves are closer to the measurements (see Figure 10). Unfortunately we can’t state firm that these points

are sure to be the real global minima, of course, but without any doubt they mean better optical model and give more precise information about the samples.

4. Conclusion

The use of Grid and the Random searches as pre-searches of gradient method seemed to be less effective in ellipsometry evaluating porous silicon samples. The SA and GA are more complex algorithms with better efficiencies and they may be powerful if the dimension of the search space is higher.

The Gradient Cube search seemed to be an effective global search algorithm in ellipsometry offering unlimited parameter space search by time, global escaping hill climbing ability due to the expanding-cube-point selection strategy (global search), and enough precision due to the involved gradient method (local search). It may useful in those cases if the optical model is complicated enough, the dimension of the search space is medium (5-7 parameters) and in consequence the error surface is hilly. If the dimension of the search space is low enough and the a-priori knowledge is precise the single use of a gradient search may also be powerful. If the dimension of the search space is too large (high number of optical model parameters) the Gradient Cube search needs long computation time, and therefore it's better to use a random-base pre-search and then a gradient search.

5. Acknowledgement

Hereby the authors would say special thanks to G. Horváth (Budapest Technical University, Hungary) for the helpful discussions and advice, and to S. Strehlke (*C.N.R.S.-Bellevue, Bellevue, France*) for making the ellipsometric samples and measurements. Partial supports from OTKA grants N° T033072 and T030441 are greatly appreciated.

References

1. Azzam, R.M.A. and Bashara N.M. (1977) *Ellipsometry and Polarized Light*, North-Holland, Amsterdam.
2. Collins R.W., Aspnes D.E. and Irene E.A. (eds) (198*), *Thin Solid Films*, Proc. of 2nd Int. Conf. on Spectroscopic Ellipsometry, 313–314 and papers therein.
3. Dekkers, A. and Aarts, E. (1991), Global Optimization and Simulated Annealing, *Mathematical Programming* 50: 367–393.
4. Fletcher R. (1971), *A Modified Marquardt Subroutine for Non-Linear Least Squares*, U.K.A.E.A. Research Group, Atomic Energy Research Establishment, Harwell.
5. Fried M., Lohner T., Polgár O., Petrik P., Vázsonyi É., Bársony I., Piel J.P. and Stehle J.L. (1996), Characterization of different porous silicon structures by spectroscopic ellipsometry, *Thin Solid Films* 276: 233.
6. Goldberg D.E. (1989), *Genetic Algorithms in Search, Optimization & Machine Learning* (p. 126) Addison-Wesley, Reading, Mass.

7. Holland J.H. (1975), *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, MI.
8. Jellison, Jr. G.E., Chisholm M.F. and Gobatkin S.M., Optical Functions of Chemical Vapor Deposited Thin-Film Silicon Determined by Spectroscopic Ellipsometry, *Appl. Phys. Lett.* 62: 348.
9. Kirkpatrick, Jr, S., Gelatt C.D. and Vecchi M. (1983), Optimization by Simulated Annealing, *Science* 220: 671–680.
10. Moré, J. (1977), The Levenberg-Marquardt Algorithm: Implementation and Theory, Numerical Analysis, (G.A. Watson, Ed.), *Lecture Notes in Math*. 630. Springer Verlag, Berlin.
11. Rich Elaine and Knight Kevin (1991), *Artificial Intelligence*, McGraw-Hill, Inc.
12. Vanhellemont, J., Maes, H.E. and De Veirman, A. (1991), Spectroscopic Ellipsometry Studies of SIMOX Structures and Correlation with Cross-Section TEM. *Vacuum* 42: 359–365.