Global optimization in inverse problem of scatterometry

Lekbir Afraites, Jérôme Hazart, Patrick Schiavone

Global optimization in inverse problem of scatterometry

Lekbir Afraites[1,2] Jerome Hazard[3] Patrick Schiavone[1,4]

[1]Laboratoire des Technologies de la Microélectronique CNRS, 17 rue des Martyrs, 38054 Grenoble, France
[2]Ecole Nationale des Sciences Appliquées, Safi, Université Cadi Ayyad, Maroc
[3]CEA-LETI-DOPT-SINA, 17 Avenue des Martyrs, 38054 Grenoble Cedex France
[4]UMI 2958 Georgia Tech-CNRS, Georgia Institute of Technology, 777 Atlantic Dr., Atlanta (GA) 30332, USA

Abstract In the current work, we consider the inverse problem in scatterometry which consists in determining the feature shape from an experimental ellipsometric signature. The reformulation of the given nonlinear identification problem was considered as a parametric optimization problem using the Least Square criterion. In this work, a design procedure for global robust optimization is developed using Kriging and global optimization approaches. Robustness is determined by Kriging model to reduce the number of real functional calculations of Least Square criterion. The technical of the global optimization methods is adopted to determine the global robust optimum of a surrogate model.

Keywords Inverse problem in scatterometry, Kriging, Global Optimization

1 Introduction

The process control in microelectronics manufacturing requires real time monitoring techniques. Among the different metrology techniques, scatterometry, based on the analysis of the light diffracted by microscale patterns using for example an ellipsometer, is well suited. The problem of computing the signature from a given structure shape, which is referred to as the direct problem, is dealt with using conventional Maxwell equations solvers, generally based on modal methods [1]. On the opposite, the inverse problem [2],[3], which allows the determination of the feature shape from an experimental signature is solved by parametric optimization problem using the Least Square objective function. This inverse problem is difficult to solve. On one hand, the problem is ill-posed, which requires for example the use of regularization methods. On the other hand, the use of traditional methods of optimization brings us back to a local optimum and the quality of the result depends on the initial point. To solve this problem, [7] developed a precomputed library in order to find the best parameters inside this library. Among the disadvantages of this method, the computing time of the direct problem is too expensive i.e. the long running times of the computer codes involved and the failure to simulate data which coincides better with experimental.

To deal with the local optimum and the dependence of the initial point, we propose a new approach based on the Kriging interpolation method and using the techniques of the global optimization. The Kriging interpolation method [4],[5], developed by Matheron and Krige [6] is based on the theory of regionalized variables. It is a stochastic interpolation, which has proven to be reliable when approximating deterministic behaviors [5]. Indeed, it attempts to obtain statistically the optimal prediction, i.e. to provide the best linear unbiased estimator. The basic premise of the Kriging interpolation method is that every unknown point can be estimated by the weighted sum of the known points. The method also provides a mechanism for estimating the interpolation error for any approximated point. Thus, the use of this interpolation method allows to create the response surface, and the global optimum of the problem is found using a global optimization algorithm.

The paper is organized as follows. In the second section, we present the principles of ellipsometric scatterometry and then talk about the direct and inverse problems. In section 4, we present the efficient global optimization (EGO) [8] algorithm sequentially samples results from an expensive calculation, does not require derivative information, uses an inexpensive surrogate obtained by techniques Kriging to search for a global optimum. In the final section, we present an application of the EGO algorithm to the simple synthetique example and to the inverse problem of ellipsometry.

2 Ellipsometric signature and inverse problem

Scatterometry is used as a generic term for several metrology methods, which may be described as a measurement technique for a quantitative evaluation of the geometrical or material properties of an object through the analysis of the light scattering from the surface under test. Since no imaging optics is used, the surface and the shape have to be reconstructed from intensity and/or polarization data detected in the far field. In our case, we use spectroscopic ellipsometry. The metrology device that measures the polarization change upon reflection by the sample is kept static whereas the incident wavelength is varying. As mentioned in the introduction, the direct problem is used to establish signatures from a given shape topography using a Maxwell solver. We use the Modal Method by Fourier Expansion to do that. This method is well adapted for the rectangular topography of the samples used in the microelectronic manufacturing which are of primary interest for us. During etching, the multi-wavelength λ_i, the direct problem gives numerically intensity $I_{\lambda_i} = (I_s(\lambda_i)^{theo}, I_c(\lambda_i)^{theo})$. Our goal is to solve the inverse problem [2, 3] which allow the determination of the feature shape from an experiment ellipsometric signature $m(\lambda_i) = (I_s(\lambda_i)^{exp}, I_c(\lambda_i)^{exp})$.

For this, we consider the objective function Least Square which can be written as a difference between the theoretically computed direct specter and the real measure:

$$J(L) = \frac{1}{2}\sum_{i=1}^{i=n}(I_{\lambda_i} - m(\lambda_i))^2, \tag{1}$$

where L is the set of optimization parameters. The objective of this work is to find the global optimum for this objective function J using the response surface obtained by Kriging techniques. For more details on the Kriging see [**?**, 4]. In our study, we have applied the Kriging techniques fo the reconstruction of the ellipsometric signatures [11]. Now, we present In the next section, the global optimization procedure.

3 Efficient Global Optimization Algorithm (EGO)

This section is inspired the work of Donald R.Jones, Matthias Schonlau and William J.Welch [8]. We give same technique developed in this paper. The idea is based on the optimization of the response surfaces constructed by Kriging model. The simplest way is to fit a surface and to find the minimum of the surface. However, if we process for this procedure, we can easily lead to a local minimum, and we have no information (idea) on the uncertain areas of the response surface given by Kriging method. It puts too much emphasis on exploiting the predictor and no emphasis on exploring points where we are uncertain. To eliminate this problem, we must put some emphasis on sampling where we are uncertain, as measured by the standard error of the predictor. To combine the search for local and global minimum and we take into account the uncertainties of the Kriging surfaces. We use a criterion based of the balances between local and global search is "$expected\ improvement''$. This concept is introduced in the literature at 1978 in [9]. The EGO is a surrogate (or meta) modeling technique, where the expensive objective function evaluation is replaced with a model that is both cheap to construct and to evaluate.

This technique uses a Kriging surrogate model to predict the values of the objective function as a few, sparsely distributed sample points $(y(x_1),...,y(x_n)$. These sample points are generally chosen by a space filling sampling method. The kriging technique is essentially a method of interpolation between known points that gives a mean prediction, $\tilde{y}(x)$, in addition to a measure of the variability of the prediction, $s(x)$, the error estimate standard. Another suitable global optimization technique is the direct method [10], which is employed to solve an auxiliary problem to find the next best place to sample for a minimum primary objective function. The secondary objective function used to solve the auxiliary problem in this application is the Expected Improvement ($E[I]$) objective function. The improvement function (I) is defined as the improvement of the current prediction, $\tilde{y}(x)$, at point x over the minimum value of the current set of samples, y_{min}, i,e :

$$I = max(y_{min} - \tilde{y}(x), 0). \tag{2}$$

The expected improvement, defined as the expectation of the improvement, is given by :

$$E[I] = (y_{min} - \tilde{y})\phi(\frac{y_{min} - \tilde{y}(x)}{s(x)}) + s(x)\Phi(\frac{y_{min} - \tilde{y}(x)}{s(x)}) \tag{3}$$

Where ϕ is the standard normal cumulative density function, and Φ is the standard normal probability density function.

The point at which the value of the expected improvement is maximized gives the best point at which to calculate the true objective function. The expected improvement is construceted to search for both local and global minima [8]. The surrogate model is then updated to include the newest sampled point, and the operation is repeated until the sampling point has not been found. An overview of the algorithm is given as follows:

Algorithm 1.

1. *An initial set of input parameters is selected.*
2. *The true objective function y is evaluated for all new members of the set.*
3. *A Kriging surrogate model is fitted to the values of the objective function.*
4. *Maximization of the expected improvement objective function criterion* $E[I]$*,*
5. *The result of the maximization (the next input parameters most likely to improve the true objective function) is added to the set.*
6. *The process repeats from step 2 until a predetermined number of iterations is reached or* $\dfrac{Max(E[I])}{y_{min}} < \varepsilon$

4 Numerical results

4.1 Synthetic numerical results

In order to validate and to explain the EGO algorithm, we are testing with a simple synthetic example. We consider the true function

$$f(x) = x.sin(x) + x.cos(x). \tag{4}$$

the objective is to find the global optimum of the true function (4) (the blue sold line in Figure 1). We consider that we have just the same points generated by this true function $(y(x_1), y(x_2), ..., y(x_7))$ (the red star in Figure 1), and we create response surface using the Kriging techniques (the black shaded line in Figure 1) associate the standard error MSE (the green line in Figure 1). Now, we apply the global optimization algorithm described

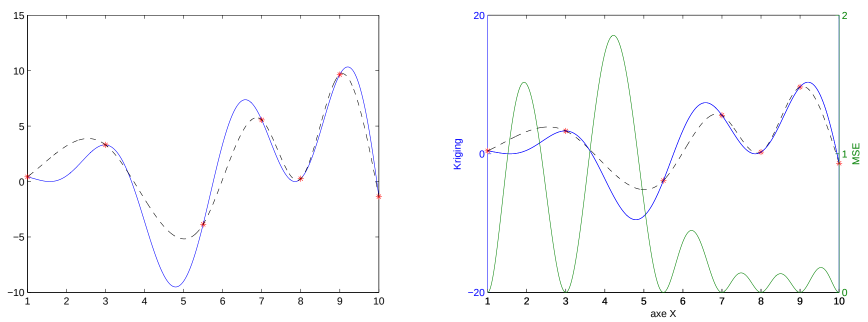

Figure 1: The true function (blue solid line), the set points used for the kriging interpolation (red star), the surface Kriging (black dashed line) and the estimator error (green solid line)

above for the problem constructed by the set the real points $(y(x_1), ..., y(x_7))$ and the surface $\tilde{y}(x)$ obtained by Kriging. In the next figure, we present the expected improvement criterion $E[I(x)]$ (the green sold line in the figure 2) and, where it's maximized, we added another point in the set, this point is evaluated by the true function 4 (the black triangle in the right Figure 2). We repeated this process, the EGO method has run using an initial sampling of 7 points to build the surrogate (the shaded line in Figure 3). A further 5 expensive function evaluation (the triangle in Figure 3) were required to find the global

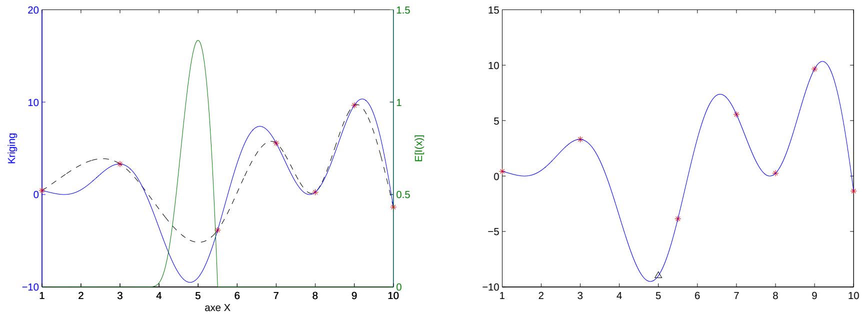

Figure 2: The true function (blue solid line), the set of points used for the kriging interpolation (red star), the Kriging surface (black dashed line), the Expected Improvement criterion (green sold line) and in the right the true function with the add point (the triangle)

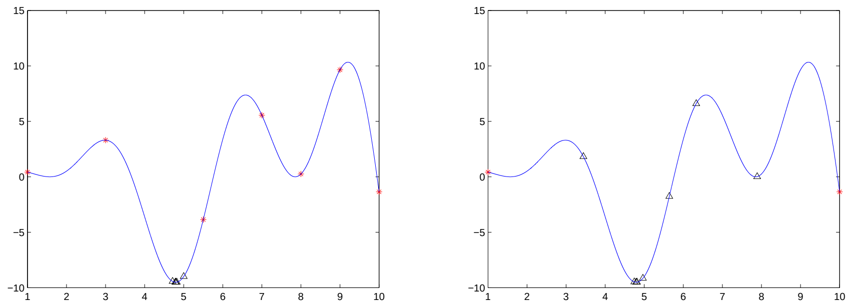

Figure 3: In the left the EGO convergence with 7 points and in the right the convergence with 2 points are shown

minimum. Now, we apply the EGO method with an initial sampling of 2 points and the EGO is able to find reasonable solution in 13 function evaluations.

4.2 Real case : a film plan of thickness

Now, we applied the global optimization algorithm combined with the Kriging model in order to find the thickness of the film plan when we know the optical indices n and k. We have the objective function described in section 2 by equation 1.

To highlight the method, we compare it with the conventional method of optimization (classical regression). The thickness approximated is $99,7393$. By the classical method based on conjugate gradient method, we found the following results :

- If the initial point is 150, the method converge to $99,4861$ with 32 number of evaluations.

- If the initial point is 170, the method converge to 318.6035 with 1474 number of evaluations.

We remark that this method is dependent on the initial point and, it's gives the local minimum.

Now, we present the results of the global optimization method. We show in the next figure, the initial points and the point which are added by this techniques. In the Figure

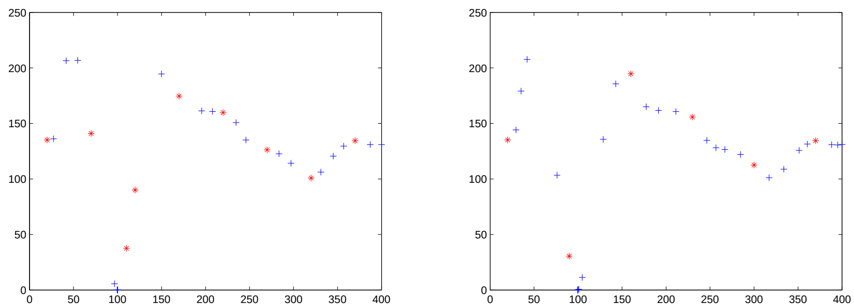

Figure 4: In the left, the EGO with 9 values for initial points, in the right, the EGO with 6 initial points

4, we initiate the algorithm by 9 points (red star in figure 4) and we have the convergence after 21 evaluations (blue plus in the figure4). By 6 points of initialization, the algorithm EGO converges after 26 evaluations. Now, we apply the EGO method with an initial

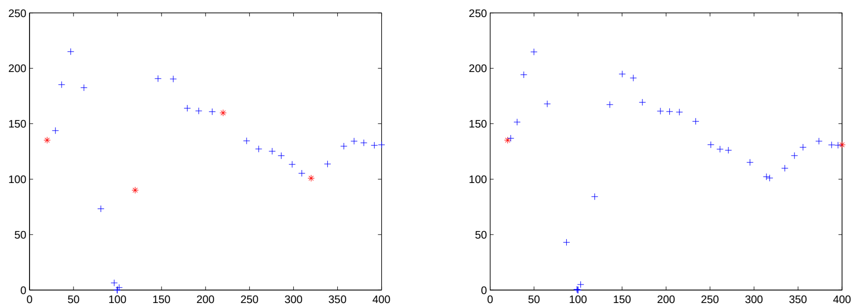

Figure 5: In the left, the EGO with 4 values for initial points, in the right, the EGO with 2 initial points

sampling of 4 and 2 points and the EGO is able to find reasonable solution in 32 and 36 function evaluations 5.

5 conclusions and future work

The advantage of this algorithm EGO, that's no dependent of the initial points and it's good mean to find the next best place to sample for a minimum objective function. The algorithm is tested on a dimension greater than one, which gives good results which will be published.

6 Bibiliography

References

[1] Richard Quintanilha *Étude du problème inverse en diffractométrie spéctroscopique* PhD thesis, INPG, Grenoble, December 2005.

[2] C.J. Raymond et al. Proceeding of the SPIE vol. 5375, pp. 564-575 (2004).

[3] V..V. Borbo, A.S. Mardezhov, A.I. Semenenko *Solution of the incorrect inverse ellipsometric problem* Proc. SPIE, Vol. 3485, 354 (1998).

[4] Sacks J, Welch W J, Mitchell T J and Wynn H P. *Design and analysis of computer of Computer experiments* Statistical Science, vol. 4(4): 409-435 (1989).

[5] Olive, M.A., Webster, R. *Kriging : a Method of Interpolation for Geographical Information Systems* Int. J. Geographic Information systems, 1990, Vol 4, N 3, PP 313-332.

[6] Cressie, N.A. *Statistics for spatial Data* Wiley, New York, 1993.

[7] Sébastien Soulan *Développement de la scatteromtrie dynamique pour le suivi en temps réel de procédés. Application la microlectronique* PhD thesis, UJF, Grenoble, December 2008.

[8] Donald R.Jones, Matthias Schonlau and William J.Welch *Efficient Globam Optimization of Expensive Black-Box Functions* Jounral of Global Optimization 13: 455-1998

[9] Mockus,J., Tiesis, V. and Zilinskas, A.(1978), *The application of Bayesian methods for seeking the extremum, in L.C.W.Dixon and Szego(eds.),* Towards Global Optimization, Vol.2, pp.17-129. Noth Holland, Amsterdam.

[10] Finkel, D. E., *Direct Optimization Algorithm User Guide, center for Research in Scientific Computer,* North Carolina State University, 2003.

[11] L. Afraites, J. Hazart, P. Schiavone, *Application of the Kriging method to the reconstruction of ellipsometric signature* Microelectronic Engineering, 2009.