International Journal of Pure and Applied Mathematics
Volume 18 No. 3 2005, 385-398

MATHEMATICAL MODELLING OF THE REFRACTIVE INDEX RECONSTRUCTION THROUGH GLOBAL OPTIMIZATION

Gleb Beliakov[1 §], Stephen Buckley[2]
[1]School of Information Technology
Deakin University
221 Burwood Hwy, Burwood, 3125, AUSTRALIA
e-mail: gleb@deakin.edu.au
[2]School of Science and Technology
Charles Sturt University
Wagga Wagga, 2678, AUSTRALIA
e-mail: sbuckley@csu.edu.au

Abstract: We examine a mathematical model of non-destructive testing of planar waveguides, based on numerical solution of a nonlinear integral equation. Such problem is ill-posed, and the method of Tikhonov regularization is applied. To minimize Tikhonov functional, and find the parameters of the waveguide, we use two new optimization methods: the cutting angle method of global optimization, and the discrete gradient method of nonsmooth local optimization. We examine how the noise in the experimental data influences the solution, and how the regularization parameter has to be chosen. We show that even with significant noise in the data, the numerical solution is of high accuracy, and the method can be used to process real experimental data.

AMS Subject Classification: 65R20, 65R32, 78A05, 65K05, 90C26
Key Words: inverse problems, global optimization, integral equations, geometrical optics

Received: January 1, 2005

§Correspondence author

1. Introduction

Planar optical devices are of great importance to optoelectronics industry. Such devices typically consist of a base made of glass, and the waveguide layer – a very thin layer of transparent material (Figure 1). If the light enters the waveguide layer at a certain angle, it will remain in the waveguide layer because of the effect of total reflection. By creating waveguide layers with specific distributions of the refractive index, one achieves various optical effects, such as focusing the light (e.g., the Luneburg lens [16]), which leads to the production of useful components.

As a result of the manufacturing process, the refractive index n of the waveguide layer is not constant (either by design, or otherwise). The function $n(z)$ which models variation of the refractive index with the depth of the layer is called the refractive index profile. Many optical properties of the waveguide can be calculated from n. Conversely, one can design waveguides with the required properties by varying and adjusting n.

During the manufacturing process, one aims to obtain waveguide layers with the specified index n. For many technological reasons, it is virtually impossible to match the specifications exactly. Therefore, an important post-fabrication task is to measure the resulting index profile and compare it with the specification. There is a number of methods of non-destructive measurement of the index [23, 22, 10, 15], most of which are based on inverting WKB (Wentzel-Kramer-Brillouin) integral and nonlinear least squares fit.

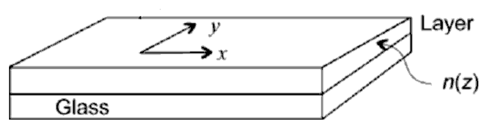

Figure 1: Structure of a planar waveguide

This paper examines a new alternative method of non-destructive index profile measurement, entirely based on geometrical optics model. We formulate a mathematical model which relates the refractive index with the paths of thin laser beams passing across the waveguide layer (Figure 2), and obtain the equations to compute the index profile from the performed measurements. We then examine the resulting ill-posed problem involving a nonlinear integral equation, and present an approach to solve it. Further we dicuss numerical solution methods, examine their performance of several model problems, and their accuracy and sensitivity to noise in the data.

2. Mathematical Model

Consider the following non-destructive waveguide diagnostics setup. A set of thin laser beams passes across the waveguide layer and exits at the opposite side (Figure 2). Since n is non-constant, the rays bend inside the layer, and exit at different angles and positions, which are recorded along with the entry angle θ. In practice, we use just one laser beam and perform measurements for different angles θ one at a time, but it is convenient to view this mathematically as a set of laser beams. We cannot observe the paths of sampling rays inside the layer, so what we have is the set of exterior measurements $\{\theta_i, h(\theta_i)\}, i = 1, \dots I$ (Figure 3). Our goal is to use these data to reconstruct the index profile $n(z)$. We assume that n does not change with x, y. We also assume $n(z)$ is continuous.

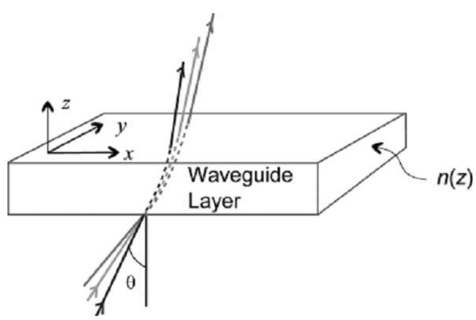

Figure 2: Experimental setup: a set of sampling rays propagating across the waveguide layer

In geometrical optics approximation, the propagation of light is expressed in the Eikonal equation

$$[\nabla S(r)]^2 = n(r)^2,$$

where $S(r)$ denotes the optical path and $n(r)$ is the refractive index in a point $r = (x, y, z)$. Since $n = n(z)$, we can write it in this form [9]

$$\frac{nx''}{1 + x'^2} + x' n_z = 0, \tag{1}$$

where $x = x(z)$ is the ray path parameterized by z. The solution to the Eikonal equation in the stratified media is given by [16]

$$p(z) = \frac{k}{\sqrt{n^2(z) - k^2}}, \tag{2}$$

where k is the integration constant which depends on θ, and $p = x'(z)$. k can be found from the initial condition

$$p(0) = \tan\theta = \frac{k}{\sqrt{n^2(0) - k^2}},$$

from which we derive $k = n(0)\sin\theta$.

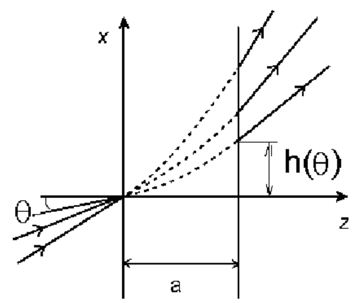

Figure 3: Sampling rays and their exit positions

Thus we can write

$$x(z) = k\int_0^z \frac{dz}{\sqrt{n^2(z) - k^2}}\,. \tag{3}$$

If the depth of the waveguide layer is a, then we can relate the entry angles and the exit positions by

$$x(a) - x(0) = h(k) = k\int_0^a \frac{dz}{\sqrt{n^2(z) - k^2}}. \tag{4}$$

Equation (4) is the main equation that will be used for index profile reconstruction. Without loss of generality we can put $a = 1$. Let us now use other available information to simplify the problem. We assume $n_0 = n(0)$ to be known (since we have access to the outer boundary at $z = 0$, we can measure $n(0)$ directly, e.g., by using liquids with matching index). We also assume the refractive index of the air and of the glass base to be known. This way we can implicitly account for the refraction at the boundaries $z = 0$ and $z = 1$ by using the law of sines, and think of $\tan\theta = \lim_{z\to 0^+} p(z), \tan\theta_1 = \lim_{z\to 1^-} p(z)$. The ratio of $\frac{\sin\theta}{\sin\theta_1}$ gives us the ratio of $\frac{n(1)}{n(0)}$. Let us express $n(z)$ in the units relative to $n(1)$ (so called effective refractive index). Consequently we have $n(1) = 1$. We also assume that $n(z)$ is a monotone decreasing function (the manufacturing process results in monotone continuous $n(z)$).

The condition that the ray entering the waveguide layer at an angle θ exits the layer at $z = 1$ is $n(z) > k$, which translates into $\theta < \sin^{-1}(n(1)/n(0)) = \sin^{-1}(1/n_0) = \theta_{max}$. The rays enter the layer at $x = 0$. For $\theta = 0$ there is no refraction and $h(0) = 0$. Because of the symmetry $h(\theta) = -h(-\theta)$. Thus we consider experimental data in the range $0 \le \theta < \theta_{max}$.

Equation (4) naturally leads to two distinct problems, the direct and the inverse. The *direct* problem consists of solving (4) for $h(k)$ with *known* $n(z)$, i.e., computing the exit position of a ray entering the media with known $n(z)$ at the origin at an angle θ. It is a variant of the ray tracing problem. We will use it later as a tool for simulating experimental data.

The *inverse* problem is the one we are interested in: to invert (4) and compute the unknown $n(z)$ given $h(k)$. We note that a similar problem arises in computing the index profile of a circular symmetric optical fibre (i.e., when $n = n(r)$). It was dealt with in detail in [11, 19], and it involves Abel type integral equations. Our case is different because equation (4) does not allow explicit inversion [17], hence we will have to solve (4) for n numerically.

3. Solution to the Integral Equation

Let us now examine in detail equation (4). This is a nonlinear integral equation with constant limits. Its solution is an ill-posed problem [20]. There is no guarantee that small changes in h will result in relatively small changes in n. Further, for the ideal case of exact measurements, the existence and uniqueness of solution of (4) follows from the fact that $h(k)$ is the solution of the direct problem (i.e., there is such $n(z)$ that generated $h(k)$ via (4)), and that equation (4) is not singular for $\theta \in [0, \theta_{max})$ and the integrand is bounded [17].

However, when we consider the case of noisy data $h(k) + \epsilon$, the existence of the solution of the inverse problem (4) is not guaranteed. It is a common situation in the inverse problems [20], where neither existence nor continuous dependence on the data do not hold.

A common approach to solution of ill-posed problems is Tikhonov regularization [20]. It consists in minimizing the functional

$$||An - h|| + \alpha||Dn||,$$

where A is the integral operator (4) acting on n, and D is a differentiation operator. It is customary to use second derivatives, which penalize unwarranted oscillations of n, in which case we solve the minimization problem

$$\min \left|\left| h(k) - k \int_0^1 \frac{dz}{\sqrt{n^2(z) - k^2}} \right|\right| + \alpha \int_0^1 [n''(z)]^2 dz. \tag{5}$$

Let us use a discretized representation of n through its values at certain points in (0,1), $x_j, j = 1, \ldots, J$, i.e., $n_j = n(x_j)$. Under our assumption of monotonicity of n, $n_0 \geq n_1 \geq \ldots \geq n_J \geq 1$. Then we have the approximation of the integral in the Gauss formula

$$\int_0^1 \frac{dz}{\sqrt{n^2(z) - k^2}} \approx \sum_{j=1}^{J} \frac{w_j}{\sqrt{n_j^2 - k^2}}.$$

The nodes x_j are chosen as zeros of Legendre polynomials, and weights w_j are then found from the standard tables [1]. Gauss integration results in a very high precision for smooth profiles, with only few nodes required ($J = 5$ or $J = 7$). The second integral in (5) is approximated analogously, using second order divided differences denoted n''_j.

The regularization parameter α is chosen according to the noise in the data. We return to this issue in the following sections. Let us now use the set of experimental measurements $\{(\theta_i, h(\theta_i))\}_{i=1}^{I}$, and convert problem (5) into

$$\text{minimize} \sum_{i=1}^{I} \left[h(k_i) - k_i \sum_{j=1}^{J} \frac{w_j}{\sqrt{n_j^2 - k_i^2}} \right]^2 + \alpha \sum_{j=1}^{J} w_j {n''_j}^2, \tag{6}$$

where as previously $k_i = n_0 \sin(\theta_i)$, and n_j are restricted by the monotonicity condition. Problem (6) will be solved numerically to obtain n.

4. Numerical Methods

Now we address the problem of numerical solution of (6). We notice that the vector of unknowns $(n_1, n_2, \ldots, n_J)$ is restricted by the conditions of monotonicity. By using a change of variables $d_j = (n_{j-1} - n_j)/(n_0 - 1), j = 1, \ldots, J + 1,$

where n_0 and $n_{J+1} = n(1) = 1$ are given, we obtain a new vector of unknowns d restricted to the unit simplex $d \in S = \{t \in R^{J+1}, t_j \geq 0, \sum_{j=1}^{J} t_j = 1\}$.

Further, the objective function in (6) is not necessarily convex, which means that problem (6) may have multiple locally optimal solutions, whereas we are interested in the globally optimal solution. The problem of multiple local minima is well known in many fields [14, 21]. For integral and differential equations it was considered in [13, 12]. There are several approaches to its solution, broadly classified into stochastic and deterministic methods. Stochastic methods (such as random search, simulated annealing) converge to the global optimum in probability, whereas deterministic methods [14] guarantee global optima. On the other hand, deterministic methods are computationally very expensive due to their slow convergence. It is customary to use a combination of various methods, like to run a deterministic global optimization method for a number of iterations, and then to rapidly improve the set of best solutions by using a local descent method.

For a relatively small number of variables (up to 10), a new deterministic global optimization method of cutting angle (CAM) has shown the potential to solve many classes of optimization problems with Lipschitz objective functions on a simplex [18, 3, 4, 5, 6]. In our case, the objective function is certainly Lipschitz (the data are registered in the interval $[0, \theta_{max})$, which translates into $\forall i, j : n_j > k_i$), and the domain is the unit simplex S, so we can apply CAM to find the global minimum of (6).

The second new optimization method we can rely on is called the discrete gradient method (DG) [2]. This is a method of local non-smooth optimization. We apply it combined with the exact penalty function approach to penalize for violation of simplex constraints. One particular feature of DG method is its ability to "skip" through shallow local minima, and converge to a deeper minimum than other local methods. We verified it experimentally in [7]. This method will be applied at the end of the optimization procedure, to rapidly improve the best solutions found by CAM. We refer the reader for the details of both methods to [18, 2, 6].

Once the optimal solution to (6) is found, we reconstruct the function $n(z)$ by interpolating the points (x_j, n_j) by a polynomial of order J, for which the Gauss integration formula is exact.

5. Simulation of Experimental Data

To verify the index reconstruction procedure, and establish the bounds of its applicability, we need to simulate the experimental data. We take some model index profile $n(z)$ and solve equation (4) with respect to $h(k)$, for various choices of k. This way we create the set of pairs $\{(\theta_i, h(\theta_i))\}, i = 1, \ldots I$, which represents the experimental data. Then, by solving (6) with these data, we reconstruct $n_j, j = 1, \ldots, J$, and fit it with a polynomial $\hat{n}(z)$. If the numerical solution is correct, then the interpolating polynomial should match the model index $\hat{n}(z) \approx n(z)$.

We can measure the accuracy of index reconstruction by using the root mean squared error and the maximal error, i.e.

$$RMSE = \left(\frac{1}{M} \sum_{m=1}^{M} (n(x_m) - \hat{n}(x_m))^2 \right)^{1/2},$$

$$MAXERR = \max |n(x_m) - \hat{n}(x_m)|,$$

for a large M and $x_m \in [0, 1]$. We shall also plot n and $\hat{n}$ one against the other.

Further, we consider the case of noisy data, i.e., we will use $h(\theta_i) + \epsilon_i$ as the experimental data, where ϵ_i are independent identically distributed random values, from a normal distribution $N(0, \sigma)$, for various noise levels σ. We will establish how sensitive our model is to the noise in the data, and how the noise can be filtered out by using regularization.

Thus let us concentrate on the solution of the direct problem: given a model index $n(z)$, solve (4) for various $k < 1$. For certain algebraic forms of $n(z)$ we were able to solve (4) exactly. Table 1 shows various models of n for which we obtained a closed form solution. The formulae for n were chosen in such a way that they model various types of dependency on z which are likely to occur in practice. They are illustrated on Figures 4-5. By adjusting parameters of each model (a, b, c), we can easily change the shape of n, and these indices appear to be flexible enough to model a wide range of behaviour. In choosing parameters of the models, we also took into account the range of variation of n that is obtained in practice (around 5%).

6. Numerical Results

We tested our method of solution of (4) on the five models from Table 1. In the case of noiseless data, numerical solution of (4) matches the model n used

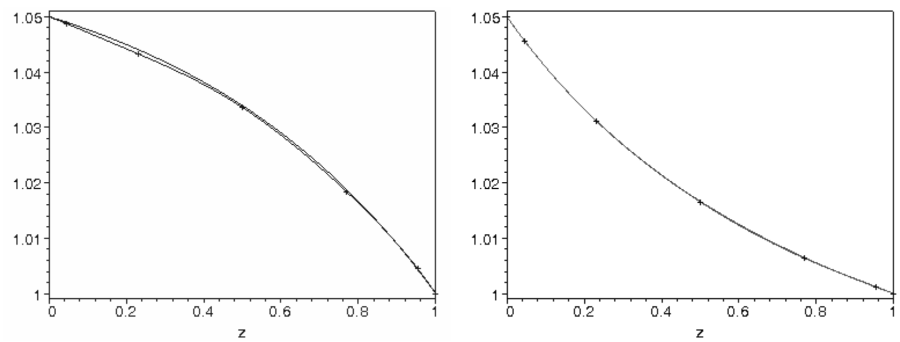

Figure 4: Reconstructed and model indices for model 2 (left) and model 3 (right). Points (x_j, n_j) are denoted by crosses

to simulate the experiment almost exactly (RMSE and the maximal error were of order of $10^{-5} - 10^{-6}$, even for a small number of experimental points $I = 20$, [8]). However, introduction of a small noise dramatically changed the picture. To compensate for noise in the data, we used a larger number of data points ($I = 100 - 500$), and a larger regularization parameter.

This strategy was successful, and even relatively large noise in the data $\sigma = 0.01$ did not prevent accurate computation of n. The reconstructed index is plotted against the model on Figures 4-5. Table 2 shows that the accuracy of 10^{-3} can be achieved in the case of noisy data. Thus we conclude that the proposed method of index profile reconstruction can be used with real experimental data with these noise levels.

An unknown with Tikhonov regularization is the regularization parameter α. Large values of α may result in oversmoothened profile, while too small values do not completely filter out the noise. For this reason we performed a study of sensitivity of our model to α. Figure 6 shows a typical plot of RMSE as the function of noise σ and regularization parameter α.

Our experiments revealed that the numerical solution of (6) is not very sensitive to large values of α. In some cases α varied three orders of magnitude with a marginal effect on the accuracy of the reconstructed index. Thus the value of α could be chosen from a large interval.

We also computed optimal values of α for various levels of noise in the data (i.e., the values leading to the smallest RMSE). They are presented in Table 2. This table can serve as a guide for choosing α when processing real experimental data: for a given noise level and the number of experimental points, we choose

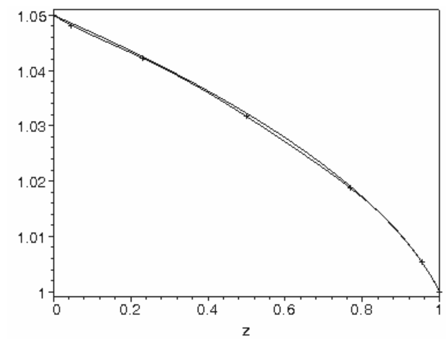

Figure 5: Reconstructed and model indices for model 4 (left) and model 5 (right). Points (x_j, n_j) are denoted by crosses

the value in Table 2. Not knowing the true $n(z)$, there is no way to establish whether such α is the optimal choice, but taking into account little sensitivity of the result to the exact value of α, such choice would be the best one.

Model	$n(z)$
1	$az + b$
2	$b - ae^{cz}$
3	$\frac{a}{1+cz} + b$
4	$c - a\sqrt{1 + bz}$
5	$b - \frac{c}{1+\exp(-a(z-0.5))}$

Table 1: Model index profiles for simulating experiment

7. Conclusion

We presented a mathematical model of the refractive index profile reconstruction, in which the unknown index profile is related to the experimental data in a nonlinear integral equation. Numerical solution of such integral equation is an ill-posed problem. We applied the method of Tikhonov regularization, and computed the index profile with high accuracy.

As a tool in numerical solution of the integral equation, we used two new methods of global and non-smooth optimization, the cutting angle and discrete

Model	Level of noise	Number of data	RMSE ($\times 10^{-3}$)	MAXERR ($\times 10^{-3}$)	Optimal α
1	0.001	50	0.027	4.25	215
	0.001	100	0.013	3.56	125
	0.001	200	0.007	5.10	130
	0.01	50	0.11	5.68	112
	0.01	100	0.043	9.26	102
	0.01	200	0.016	7.29	63
2	0.001	50	0.35	5.54	0.021
	0.001	100	0.21	4.10	0.04
	0.001	200	0.20	1.96	0.03
	0.01	50	3.77	10.1	0.006
	0.01	100	1.14	8.35	0.035
	0.01	200	0.95	9.35	0.013
3	0.001	50	0.36	9.84	0.3
	0.001	100	0.066	4.68	0.033
	0.001	200	0.082	3.99	0.016
	0.01	50	0.50	9.26	0.31
	0.01	100	0.58	8.50	0.023
	0.01	200	0.58	10.1	0.02
4	0.001	50	0.47	4.65	0.017
	0.001	100	0.36	3.64	0.03
	0.001	200	0.37	3.61	0.02
	0.01	50	4.1	9.64	0.004
	0.01	100	1.12	7.49	0.02
	0.01	200	0.96	9.52	0.008
5	0.001	50	0.39	4.15	0.019
	0.001	100	0.25	5.04	0.1
	0.001	200	0.20	4.96	0.082
	0.01	50	2.3	7.63	0.09
	0.01	100	1.3	6.18	0.015
	0.01	200	0.98	5.55	0.011

Table 2: Accuracy of index reconstruction as a function of noise and the number of observations, and the optimal regularization parameter α

Figure 6: RMSE as a function of the noise in the data and regularization parameter, for model 4. Minimizing such function allows one to find the optimal regularization parameter for a given level of noise.

gradient methods. These methods have allowed us to find the best solution in the presence of multiple local optima.

We also performed a study of the accuracy of the solution as a function of the noise in the data, the number of data and the regularization parameter. We found optimal values of the regularization parameter for given noise and number of data points. These values give guidance for choosing regularization parameter when processing real experimental data.

References

[1] M. Abramowitz, A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, National Bureau of Standards, Washington (1981).

[2] A. Bagirov, A method for minimization of quasidifferentiable functions, *Optimization Methods and Software*, **17** (2002), 31-60.

[3] A. Bagirov, A. Rubinov, Global minimization of increasing positively homogeneous function over the unit simplex, *Annals of Operations Research*, **98** (2000), 171-187.

[4] L.M. Batten, G. Beliakov, Fast algorithm for the cutting angle method of global optimization, *Journal of Global Optimization*, **24** (2002), 149-161.

[5] G. Beliakov, Geometry and combinatorics of the cutting angle method, *Optimization*, **52** (2003), 379-394.

[6] G. Beliakov, The cutting angle method – a tool for constrained global optimization, *Optimization Methods and Software*, **19** (2004), 137-151.

[7] G. Beliakov, A. Bagirov, J. E. Monsalve, Parallelization of the discrete gradient method of non-smooth optimization and its applications, In: *Proc. of the Third Intl. Conf. on Computational Science*, Volume **3** of *Springer-Verlag Lecture Notes in Computer Science (LNCS)*, Springer-Verlag, Heidelberg (2003), 592-601.

[8] G. Beliakov, S. Buckley, Reconstruction of the refractive index profile of planar waveguides using ray tracing analysis, In: *Conf. on the Optical Internet and Australian Conf. on Optical Fibre Technology*, Melbourne (2003), 384-387.

[9] G. Beliakov, D.Y.C. Chan, Analysis of inhomogeneous optical systems by the use of ray tracing .1. planar systems, *Applied Optics*, **36** (1997), 5303-5309.

[10] K. S. Chiang, C.H. Wong, S.Y. Cheng, H.P. Chan, Refractive-index profiling of graded-index planar waveguides from effective indexes measured with different external refractive indexes, *J. of Lightwave Technology*, **18** (2000), 1412-1417.

[11] P.L. Chu, Nondestructive measurement of index profile of an optical-fibre preform, *Electron. Lett.*, **13** (1977), 736-738.

[12] E.A. Galperin, E.J. Kansa, Application of global optimization and radial basis functions to numerical solutions of weakly singular volterra integral equations, *Computers Math. Applic.*, **43** (2002), 491-499.

[13] E.A. Galperin, Q. Zheng, Solution and control of the pde via global optimization methods, *Computers Math. Applic.*, **25** (1993), 103-118.

[14] R. Horst, P. Pardalos, N.V. Thoai, editors, *Introduction to Global Optimization*, Second Edition, Volume **48** of *Nonconvex Optimization and its Applications*, Kluwer Academic Publishers, Dordrecht (2000).

[15] J.M. Kubica, Noise limits in reconstruction of optical parameters of silica-based double-layer planar waveguides on silicon, *J. of Lightwave Technology*, **20** (2002), 114-119.

[16] R.K. Luneburg, *Mathematical Theory of Optics*, University of California Press, Berkeley (1964).

[17] A.D. Polyanin, A.V. Manzhirov, *Handbook of Integral Equations*, CRC Press, Boca Raton (1998).

[18] A.M. Rubinov, *Abstract Convexity and Global Optimization*, Volume **44** of *Nonconvex Optimization and its Applications*, Kluwer Academic Publishers, Dordrecht, Boston (2000).

[19] C. Saekeang, P.L. Chu, Nondestructive determination of refractive index profile of an optical fiber: backward light scattering method, *Applied Optics*, **18** (1979), 1110-1116.

[20] A. Tikhonov, V. Arsenin, *Solutions of Ill-Posed Problems*, Winston and Sons, Washington (1977).

[21] A. Törn, A. Zilinskas, *Global Optimization*, Springer-Verlag, Berlin; New York (1989).

[22] H.P. Uranus, M.O. Tjia, Refractive index profile reconstruction of planar waveguides from its mode intensity using characteristic matrix mehod, *J. Nonlinear Optical Physics and Material*, **10** (2001), 169-180.

[23] J.M. White, P.F. Heidrich, Optical waveguide refractive index profiles determined from measurement of mode indices: a simple analysis, *Applied Optics*, **15** (1976), 151-155.