Thickness gauging of thin layers by laser ultrasonics and neural network

F Lefèvre, F Jenot, M Ouaftouh, M Duquennoy, M Ourak

Institut d'Electronique de Microélectronique et de Nanotechnologie Dēpartement Opto-Acousto-Electronique (UMR CNRS 8520), Université de Valenciennes, Le Mont Houy, 59313 Valenciennes Cedex 09, France.

Fabien.lefevre@univ-valenciennes.fr

Abstract. Non destructive testing has been performed on a thin indium layer deposited on a two inches silicon wafer. Guided waves were generated and studied using a laser ultrasonic setup, and a two-dimensional Fourier transform technique was employed to obtain the dispersion curves. The inverse problem, in other words the determination of the layer thickness and the elastic constants of the substrate, has been solved by means of a feedforward neural network. These parameters were then evaluated simultaneously, the dispersion curves being entirely fitted. The experimental results show a good agreement with the theoretical model. This inversion method was found to be prompt and easy to automate.

1. Introduction
Silicon wafers are widely used in microelectronics as substrates to receive microelectronics devices. In recent years, growing interest and concern for thin layer deposition has been observed. Numerous methods have been developed to characterize such films using, for example, nanoindentation [1], x-ray reflectivity [2] or laser ultrasonics [3,4].

In this paper, guided waves are probed to study an indium layer deposited on a silicon substrate. Experimental dispersion curves are obtained in a non contact and non destructive way, and the inverse problem is solved by a feed-forward neural network [5]. The experimental dispersion curves are fitted on several modes which allows us to find out simultaneously the layer thickness and the elastic characteristics of the silicon substrate.

2. Lamb wave propagation in anisotropic bilayered structures
The transfer matrix method [6] has been used to describe the propagation of guided waves in a bilayered platelike structure. Such a structure is depicted in figure 1, where the exponents (1) and (2) refer to the different layers, whereas the indices 1, 2, and 3 refer to the different directions.

In this method, for each layer, the displacements and stresses at the same surface are expressed as a vector $P$. The relationship between the $P$, vector at the top layer surface and the $P$, vector at the bottom surface of the same layer is given by:

$$P_{\perp}^{(k)} = A^{(k)} \cdot P_{\perp}^{(k)}$$  \hspace{1cm} \text{(1)}$$

where $A^{(k)}$ is the transfer matrix of the anisotropic layer $k$. The use of continuity conditions for the displacements and the stresses at the interface between the two layers leads to the expression of the
global transfer matrix $A$. This matrix links the surface vector $P_{1}^{(1)}$ of the first layer to the surface vector $P_{2}^{(2)}$ of the second layer:

$$A = A^{(1)}, A^{(2)}$$

(2)

Using free boundary conditions to $P_{1}^{(1)}$ and $P_{2}^{(2)}$ allows us to obtain an equation which solutions can exist if and only if:

$$\begin{vmatrix}
A_{11} & A_{12} & A_{13} \\
A_{21} & A_{22} & A_{23} \\
A_{31} & A_{32} & A_{33}
\end{vmatrix} = 0$$

(3)

This problem has been solved in the case of a 4 µm thick indium layer deposited on a 2 inches silicon wafer. The formulation employed to express the components of the different transfer matrices was the one of Nayfeh [7]. The silicon having a diamond cubic structure, which implies an anisotropic behaviour, the resolution has been made in the (100) direction. The results are presented into dispersion curves in figure 2.

Figure 1. The coordinate systems of a plate composed of two anisotropic layers of thickness $d^{(1)}$ and $d^{(2)}$.

Figure 2. Theoretical dispersion curves for a 275 µm thick silicon wafer coated with a 4 µm thick indium layer.

3. Experimental setup

The system used to generate the acoustic waves is schematically shown in Figure 3. A 10 ns duration Q-switched Nd-YAG laser pulse of 532 nm wavelength was focused at the sample surface as a line source of about 5 mm length and 0.5 mm width [8]. The energy per pulse was around 6 mJ, which allowed us to work in the thermoelastic mode.

The normal displacement of each Lamb mode was detected by a Mach-Zehnder type interferometer with a power of 100 mW and a large bandwidth (200 kHz to 45 MHz). The received signals were sampled and averaged by a digital oscilloscope before acquisition. Each recorded signal corresponded to an average of sixteen laser shots in order to improve the signal-to-noise ratio. Motorized motion tables allowed us to move the laser line source to perform measurements at different distances between source and detector.

Figure 3. Experimental setup for the laser generation and detection of the guided modes.
4. Neural networks

The sample’s dispersion curves are characterized by several modes. Classical fitting methods are consequently difficult to implement and time consuming in this case. A more global method, using feed-forward neural networks, has been investigated.

Feed-forward neural networks (FF networks) are widely used to approximate functions. They are known by many different names, such as “multilayer perceptrons” [9]. Figure 4 illustrates a one-hidden-layer FF network with inputs $X_1, \ldots, X_n$ and output $Y$. Each arrow in the figure symbolizes a parameter in the network which is divided into layers. The input layer consists of just the inputs to the network. Then follows a hidden layer, which consists of any number of neurons or hidden units placed in parallel. Each neuron performs a weighted summation of the inputs, which then passes a nonlinear activation function $\sigma$, also called the neuron function.

Mathematically the functionality of a hidden neuron can be described by $F$, such as:

$$F = \sigma \left( \sum_{j=1}^{n} X_j \cdot w_j + b_j \right)$$  \hspace{1cm} (4)

where the weights $w_j, b_j$ are symbolized with the arrows feeding into the neuron.

The network output is formed by another weighted summation of the outputs of the neurons in the hidden layer. This summation on the output is called the output layer. The neurons in the hidden layer of the network in figure 4 are similar in structure to those of the perceptron with the exception that their activation functions can be any differential function. The output of this network is given by

$$Y = \sum_{j=1}^{n h} w_j^2 \cdot \sigma \left( \sum_{i=1}^{n} X_i \cdot w_{i,j}^i + b_{j,i}^i \right) + b^2$$  \hspace{1cm} (5)

where $n$ is the number of inputs and $n h$ is the number of neurons in the hidden layer. The variables $\{w_{i,j}, b_{j,i}, w_j^2, b^2\}$ are the parameters of the network model. The size of the input and output layers are defined by the number of inputs and outputs of the network respectively and, therefore, only the number of hidden neurons has to be specified when the network is defined.

Supervised learning can then be used to modify and to fix the parameters of the defined model. The network is fed with a training data, which consists of pairs of input values $x$, and desired outputs $y$. The weights are adjusted incrementally until the data satisfy the desired mapping as well as possible; that is, until $Y$, the network output, matches the desired output $y$ as closely as possible up to a maximum number of iterations.

The FF network in figure 4 is just one of the possible architectures. The architecture can be modified in various ways by changing the options.

**Figure 4.** A feed-forward network with one hidden layer and one output.

**Figure 5.** Neural network fit of the 2DFFT results for an indium layer deposited on a 275 $\mu$m thick silicon wafer.
5. Determination of the indium layer thickness and of the substrate’s elastic properties

Measurements were achieved with the experimental setup described earlier on a two inches silicon wafer on which 4 µm of indium has been deposited. 500 signals were recorded, the distance between the emitter and the receiver being increased of 20 µm before each new recording over a total distance of 10 mm. The sampling frequency was set to 500 MHz.

A two-dimensional spectral analysis was applied to the time signals and the results are presented as contour plots of amplitude versus phase velocity and frequency in figure 5. Four two-layer feed-forward neural networks were used to estimate the structure characteristics. They were trained using the Levenberg–Marquardt algorithm [10] and with a database of 2365 theoretical dispersion curves. The input vector of the different networks is composed of 29 velocity values, the hidden layer is made of 250 neurons, and the output layer gives the indium layer thickness and the elastic characteristics of the anisotropic substrate. The thickness and density of the substrate were assumed to be respectively 275 µm and 2330 kg/m^3. Concerning the layer, the Young modulus was set to 11 GPa, the Poisson ratio to 0.45 and the density to 7310 kg/m^3.

The inversion results are given in table 1 for the four networks. The layer thickness was moreover measured by means of a high contact profilometer. Concerning the elastic properties of the substrate, the inversion results were compared to the elastic characteristics currently used in the literature. A good agreement between the inversion solutions and the structure characteristics is observed.

Moreover, the fit of the 2DFFT results in figure 5 is found to be satisfying for the different propagating modes.

Table 1. Results of the network inversion for the indium layer deposited on a 2 in. silicon wafer.

|                     | Layer thickness (µm) | C_{11} (GPa) | C_{12} (GPa) | C_{44} (GPa) |
|---------------------|----------------------|--------------|--------------|--------------|
| Structure characteristics | 4.11                | 165.6        | 63.9         | 79.5         |
| Inversion solution n°1 | 4.35                | 165.0        | 67.7         | 83.5         |
| Inversion solution n°2 | 4.03                | 164.7        | 65.3         | 81.3         |
| Inversion solution n°3 | 4.14                | 166.5        | 64.5         | 81.3         |
| Inversion solution n°4 | 4.12                | 164.6        | 65.0         | 80.9         |
| Mean of the inversion solutions | 4.16               | 165.2        | 65.6         | 81.7         |
| Inversion error in % | 1.2                  | 0.25         | 2.7          | 2.8          |

Conclusion

A laser ultrasonic method for the analysis of guided waves in an anisotropic bilayered structure has been presented. Dispersion curves, with a clear identification of multimode guided waves, were obtained with a 2DFFT. Determination of the layer thickness and of the substrate elastic parameters has been made in an original way using feed-forward neural networks.

References

[1] Chudoba T, Schwarzer N, Linss V, and Richter F 2004 Thin Solid Films 239 469–470.
[2] Santucci S, la Cecilia A V, Digiacomo A, Phani R A and Lozzi L 2001 J. Non- Cryst. Solids 280 228.
[3] Bonello B, Perrin B, Romatet E and Jeannet J C 1997 Ultrasonics 35 223.
[4] Reverdy F and Audoin B 2001 J. Appl. Phys. 90 4829.
[5] Lefèvre F, Jenot F, Ouafouh M, Duquennoy M and Ourak 2009 Rev. Sci. Inst. 80 014901.
[6] Haskell N A 1953 Bull. Seismol. Soc. Am. 43 17.
[7] Nayfeh A H 1991 J. Acoust. Soc. Am. 89 1521.
[8] Jenot F, Ouafouh M, Duquennoy M and Ourak M 2005 J. Appl. Phys. 97 094905.
[9] Rosenblatt F 1961 Principles of Neurodynamics, Spartan Washington D.C.
[10] Marquardt D 1963 SIAM J. Appl. Math. 11 431.