Measurement 177 (2021) 109238

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Nonparametric estimation of SiC film residual stress from the wafer surface profile

Olga Savchuk [a,*], Alex A. Volinsky [b,*]

[a] Department of Mathematics and Statistics, University of South Florida, Tampa, FL 33620, USA
[b] Department of Mechanical Engineering, University of South Florida, Tampa, FL 33620, USA

ARTICLE INFO

Keywords:
SiC
Residual stress
Substrate profile
Least squares regression
Nonparametric regression
Second derivative estimation
Differenced method
LowLSR

ABSTRACT

Thin film residual stress is proportional to substrate curvature change after film deposition, based on Stoney's equation. Curvature is approximately equal to the second derivative of substrate deflection. It is common to apply Stoney's equation locally, where the residual stress at a selected point and direction is estimated from the local substrate curvature. The locally weighted least squares regression method (LowLSR) is adapted for estimating the substrate curvature in the radial directions across the wafer from the corresponding deflection profiles measured by a profilometer. LowLSR is implemented in the R package *npregderiv* developed by the authors. The changing film thickness profiles in the radial directions are estimated using the local linear estimator and plugged into Stoney's equation. Thus, this research is the first attempt of using nonparametric statistical methods to estimate the residual stress in SiC films with non-uniform thickness.

1. Introduction

The film residual stress may cause significant film deformation and even failure. According to Stoney's equation [1], the thin film residual stress σ_f has the following form:

$$\sigma_f = \frac{E_s}{6(1-\upsilon_s)}\frac{{t_s}^2\kappa}{t_f}, \tag{1}$$

where κ is the substrate curvature after film deposition, t_s and t_f are the substrate and film thicknesses, respectively, E_s and ν_s are Young's modulus and Poisson's ratio of the substrate, correspondingly. The assumptions underlying Stoney's equation (1) are listed in [2]. In particular, it is originally assumed that the film thickness t_f and curvature κ are both spatially constant across the substrate.

Stoney's equation (1) implies the spatially uniform residual stress across the substrate, which is usually not the case in practice (see [3]). The residual stress at a point on the substrate is a function of the point's coordinates and the chosen direction in the xy-plane. This motivates using Stoney's equation in a local manner by relating the residual stress at each point and in the selected direction to the corresponding value of the curvature.

Taking into account that the blank substrates are not ideally flat when manufactured and that the substrate deflection in the z direction is small compared to its dimensions on the xy-plane, the substrate curvature change along the x coordinate after the thin film deposition can be found as

$$\Delta\kappa(x) = \frac{d^2w_2(x)}{dx^2} - \frac{d^2w_1(x)}{dx^2}, \tag{2}$$

where w_1 and w_2 are the substrate deflections before and after the film deposition, respectively, in the direction of the x-axis. It is then relevant to replace the constant substrate curvature κ in (1) with the curvature change $\Delta\kappa(x)$ computed according to (2).

The substrate curvature is commonly estimated by first finding the polynomial least squares regression (LSR) estimate of the deflection function and then taking the second derivative of the obtained fit (see [4]). To comply with the assumptions underlying Stoney's equation (1), the second-order polynomial regression method is typically used, thus, producing the constant curvature estimate along x. An obvious drawback of this approach is disregarding a well-known fact that differentiating an optimally found regression fit does not produce an optimal estimate of the corresponding regression function's derivative (see [5]). In addition, the polynomial regression method does not have enough flexibility for checking the assumption of the constant substrate curvature and capturing the changing curvature pattern along x.

Because of the disadvantages of the LSR method, we turned to the

* Corresponding authors.
E-mail addresses: osavchuk@usf.edu (O. Savchuk), volinsky@usf.edu (A.A. Volinsky).

https://doi.org/10.1016/j.measurement.2021.109238
Received 25 January 2021; Received in revised form 22 February 2021; Accepted 1 March 2021
Available online 10 March 2021

nonparametric approach for estimating the second derivative of a deflection function. In this approach, the data are not squeezed into some predetermined, often inappropriate, model but are rather allowed to speak for themselves. This, generally, might lead to disclosing certain unexpected hidden patterns and features of a function being estimated. In our application, using nonparametric methods for estimating the second derivatives of a deflection function allows for obtaining the curvature profile along the x-axis and, thus, assessing the assumption of the constant curvature. Furthermore, plugging the obtained profiles into Stoney's equation (1) allows for a more realistic pointwise assessment of the thin film residual stress at a given point in the direction of the x-axis.

In this research, we estimated the second derivatives of the deflection functions using the locally weighted least squares regression (LowLSR) method of Wang and Lin [6]. This is a nonparametric method of the so-called differenced type where the second derivative is estimated by first using the data to compute a sequence of the specific difference quotients and then using the locally weighted LSR. The LowLSR method is implemented in the R package *npregderiv* [7] that accompanies this article. Some alternative frequently used nonparametric methods for estimating the second derivative include the smoothing splines (see [8]), the local polynomial regression (LPR) method (see [9,10]), and the difference-based methods of De Brabanter et al. [11], Liu and De Brabanter [12], Dai et al. [13], and Wang et al. [14]. It is worth mentioning that the recently developed method [14], referred to as LowLAD, is similar in spirit to LowLSR. The difference is that LowLAD uses the least absolute deviation (LAD) instead of LSR. In the Appendix, we show that LowLSR outperforms LowLAD in the case of normally distributed error terms, which is a frequently used assumption in practice. Another reason for preferring LowLSR, is because the method is shown to produce better results compared to its frequently used competitors (LPR and the method of penalized smoothing splines) in the simulation study [6].

The film thickness may vary across the substrate. For a more accurate pointwise assessment of the film residual stress in the x-direction, it is reasonable to replace the constant film thickness in (1) with the film thickness function $t_f(x)$. In this article, we used the local linear estimator (LLE) [15] to estimate $t_f(x)$ for each of the considered scan orientations from the corresponding thickness measurements taken at the selected points along the x-axis. The LLE is a nonparametric method that is frequently used for estimating the regression functions.

While contact profilometers have been previously employed to measure the wafer profiles (see [4,16,17]), this article is the first attempt of using the nonparametric statistical methods to estimate the residual stress in SiC films with non-uniform thickness, which is the major contribution of this research. The accompanying R package *npregderiv* [7] makes our research results (the code and the datasets) reproducible and available to others to solve similar or related problems.

The rest of the article is organized as follows. The materials and methods used in the experiment are discussed in Section 2. The results of the data statistical analysis are provided in Section 3, where estimates of the residual stress profiles are displayed. The obtained results are further discussed along with the directions of future research in Section 4. Section 5 outlines the main conclusions. The Appendix proves that LowLSR is more efficient than LowLAD in the case of the evenly spaced design and the normally distributed random errors, which is a common practical setting.

2. Materials and methods

2.1. Experiment

The experiment involved several (1 0 0) Si wafers with a 50 mm diameter. The 3C-SiC films were deposited on the substrates by chemical vapor deposition (CVD). Fig. 1(a) shows one of the wafers. The film exhibits the interference fringes under an optical flat that are explained by the film thickness non-uniformity and wafer curvature.

2.1.1. Substrate deflection measurements

The substrate deflection measurements were performed by using a Tencor P-20H profilometer in the 0° and 90° radial directions, as explained in Fig. 1(b). The profilometer uses a mechanical stylus that takes measurements with a vertical (height) resolution of 1 nm. The deflection measurements were recorded at the equidistant points separated by 5 μm according to the coordinate system shown in Fig. 1(b), where for each direction the x-axis passes through the wafer center. In the case of 0°, the origin of the coordinate system is placed at the wafer side opposite the wafer flat. In the 90° case, the measurements are recorded from one wafer's side to another as shown in Fig. 1(b).

In fact, for each direction, we measured deflection twice: before and after the thin film deposition. At each measurement run, the wafer was placed on three steel balls, and the wafer flat and side reference points were used to make sure that each time the wafer is placed in the same position. The tool is equipped with a microscope and the x-y stage with

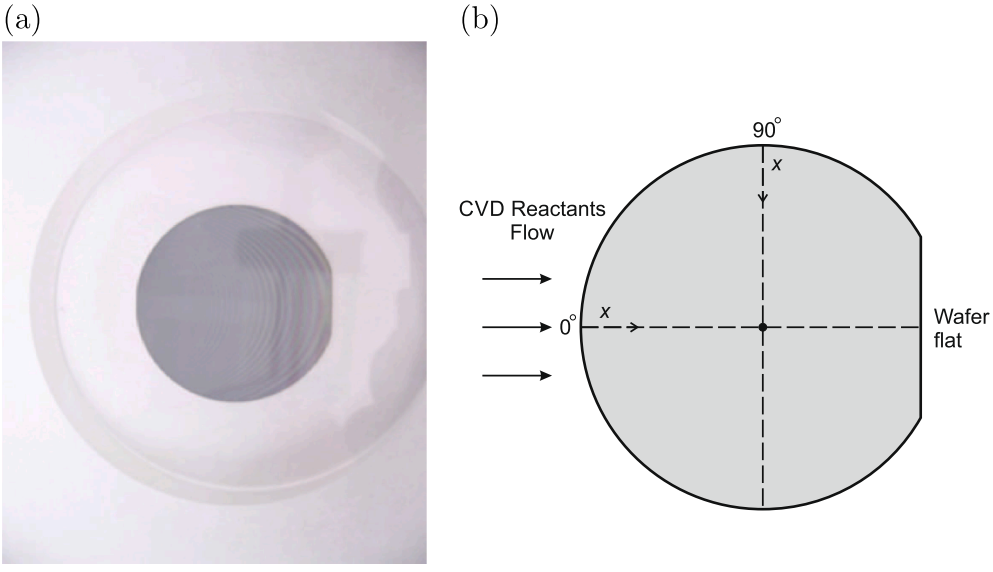


Fig. 1. (a) Photograph of the 100 mm optical flat on top of the 50 mm (1 0 0) Si wafer with the 3C-SiC film. (b) Diagram of the substrate with different scan orientation angles.

the sub-micron position accuracy.

Since the long scan surface profilometer can be damaged if the stylus runs off the wafer, the length of the scanned interval in each setting was smaller than the wafer's diameter. This resulted in n = 7755 data points for each run of measurements. Notice that our measurement details do not interfere with the described algorithms of the statistical data analysis that this article is intended to outline.

2.1.2. Film thickness measurements

The thickness of the deposited 3C-SiC films was measured by the Fourier transform infrared (FTIR) spectrometer with respect to the coordinate system shown in Fig. 1(b) at the equidistant points separated by 5 mm along the 0° and 90° scan orientations. The thickness measurement resolution is in the order of nanometers. The position deviation during the measurement run is in the order of a micron.

2.2. Statistical methods for estimating the second derivative of a deflection function based on the deflection measurements

We adopt the ordinary regression model that assumes that the data points Y_i, $i = 1, \cdots, n$, are taken at the design points x_i, generated as

$$Y_i = r(x_i) + \varepsilon_i, i = 1, \cdots, n, \tag{3}$$

where r is a regression function, and $\varepsilon_i, i = 1, \cdots, n$, are independent and identically distributed error terms that have probability density f with the mean $E(\varepsilon_i) = 0$ and the variance $Var(\varepsilon_i) = \sigma^2 < \infty$. We are interested in estimating $r^{(2)}(x_i)$, the second derivative of r at x_i.

In our case, Y_i (in μm) are the deflection measurements, $n = 7755$, the regression function r in (3) corresponds to a deflection function w, and the design points are generated as

$$x_i = \frac{5(i-1)}{n} (\text{in } \mu\text{m}) \tag{4}$$

That is, the design data are evenly spaced. In this application, one needs to estimate the second derivatives of $w_1(x)$ and $w_2(x)$, the deflection functions before and after the deposition process. For convenience, the subscripts are dropped in describing the statistical methods below.

As noted above, the deflection measurements are not taken around the edges of the substrate. For convenience, in each considered setting, the data were shifted horizontally such that $x_1 = 0$. Following [6], it is assumed that w is at least four times continuously differentiable over the estimation interval.

2.2.1. Polynomial LSR

Most often, the second derivative $\frac{d^2w(x)}{dx^2}$ is estimated from the wafer profile data using the polynomial least-squares regression method. First, the deflection function $w(x)$ is estimated on its own either on the whole estimation interval or piecewise by preliminary dividing the interval into the segments [18]. Then, the second derivative of the obtained fit is calculated and used as the second derivative estimator of $w(x)$. The quadratic regression method is usually used to comply with the assumption of the constant curvature underlying Stoney's equation (1).

The major disadvantage of using polynomial regression in the described way is that the obtained second derivative estimator is not optimal in any way because of the reason mentioned in the introduction. Using the quadratic regression method and obtaining the constant curvature estimate may lead to overlooking the changing curvature pattern. When using polynomial regression of order higher than two, one faces the problem of choosing the polynomial order. Inappropriately selected polynomial order may result in an inaccurate estimation of some important features of the deflection function, such as its modes and/or inflection points. Thus, the too-low degree of a polynomial may lead to smoothing out the modes, while the too-high degree may result in some fake modes and inflection points in the estimate. The highest degree of a polynomial obtained after differentiation determines the overall behavior of the second derivative estimate, especially in the boundary regions. Estimation errors are magnified after computing the second derivative of a deflection function. The aforementioned problems are illustrated in the simulation example in Section 2.2.2.

Using polynomial regression in segments does not help to overcome the problems mentioned above and, additionally, artificially produces discontinuities in the second derivative estimate in the boundary points of the successive intervals. Moreover, in the segmentation approach, the researcher needs to decide on the number of segments and locations of the boundary points. Subjective choices of these smoothing parameters may lead to radically different estimation results and conclusions reached by different researchers.

2.2.2. LowLSR

In this research, we adopted the LowLSR method of Wang and Lin [6] for nonparametric estimation of the second derivative of a deflection function $w(x)$. This differenced method is developed for the evenly spaced design settings, which is the case of described experiments. It targets the second derivative directly without preliminary estimating the regression function and then differentiating the fit.

The method's algorithm under model (3) requires computing the following difference quotients:

$$Y_{ij}^{(2)} = \frac{Y_{i-j} - 2Y_i + Y_{i+j}}{j^2/n^2}, \ k+1 \le i \le n-k, \ 1 \le j \le k,$$

where k is the smoothing parameter of the method that determines the distance between observations in the computed differences. To estimate the second derivative of a regression function r at the design point x_i, the asymptotically optimal k has the following form:

$$k_{opt,i} = C \left(\frac{1}{r^{(6)}(x_i)^2} \right)^{1/13} n^{12/13}, \tag{5}$$

where C is a constant that depends on the parameters of the model (3) (see [6] for details). The above formula is useful in the case $r^{(6)}(x_i) \neq 0$. Notice that in this research it is only assumed that the fourth derivative of $w(x)$ exists. Even if this assumption is extended, the authors are not willing to assume that the sixth derivative of $w(x)$ is nonzero at all x_i. This makes expression (5) inapplicable in this case. In either case, (5) has limited practical use since it involves the sixth derivative of the underlying regression function that is difficult to estimate. Even the authors of [6] do not rely on (5), but rather implement their method based on the values of k ranging from $0.02n$ to $0.2n$. Their estimation results for different k are fairly consistent (see Fig. 2 in [6]) with somewhat smoother estimates corresponding to the larger values of k.

The algorithm for estimating $r^{(2)}(x_i)$, the second derivative of the regression function at x_i, in the interior region, that is for $k+1 \le i \le n-k$, is described in [6] and outlined below. Compute the following matrices:

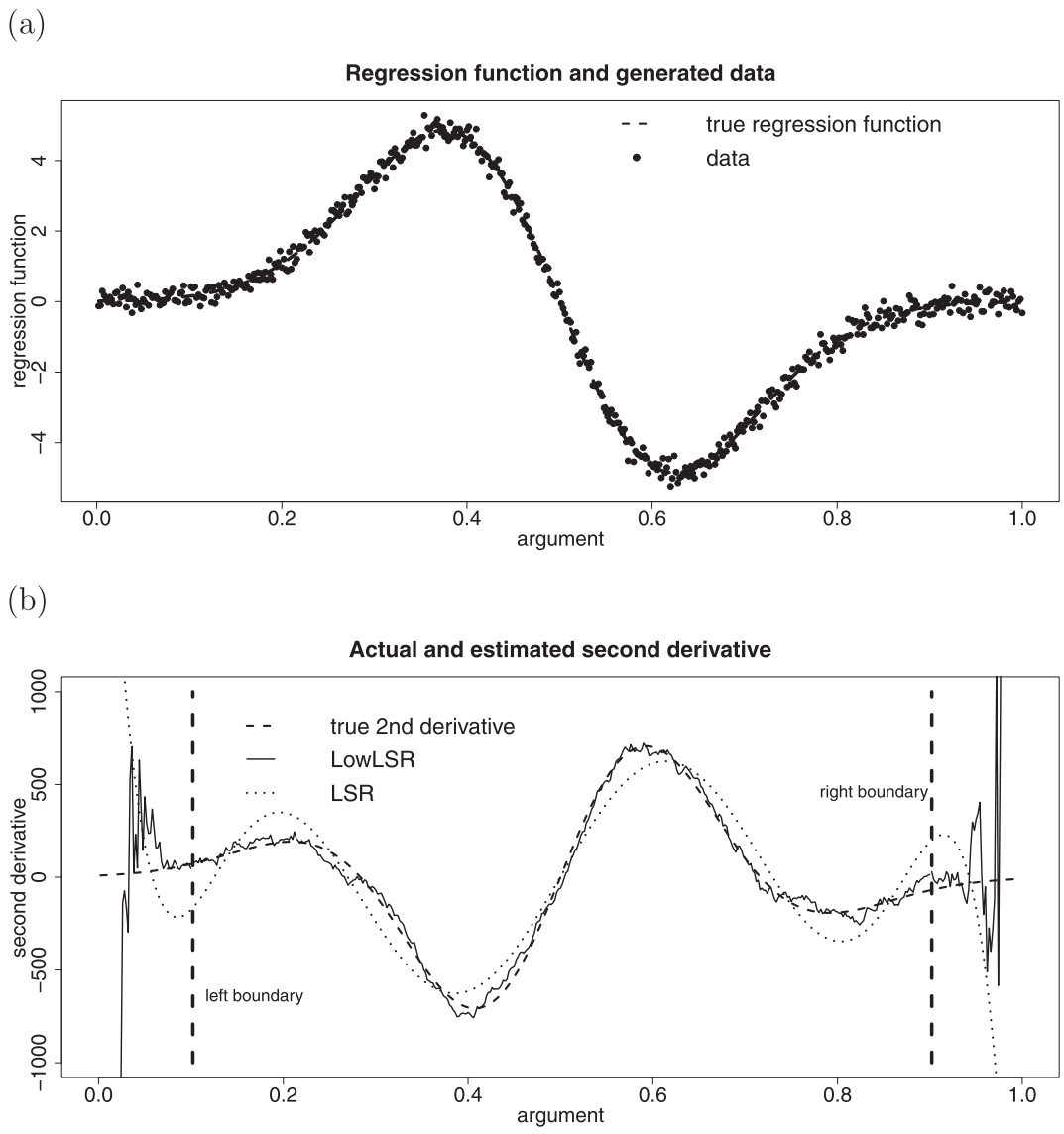


Fig. 2. (a) True regression function and generated data; (b) True second derivative and the LowLSR and LSR estimates.

$$D=\begin{pmatrix} 1 & 1^2/n^2 & n^2/1^2 \\ 1 & 2^2/n^2 & n^2/2^2 \\ \vdots & \vdots & \vdots \\ 1 & k^2/n^2 & n^2/k^2 \end{pmatrix}, Y_i^{(2)}=\begin{pmatrix} Y_{i1}^{(2)} \\ Y_{i2}^{(2)} \\ \vdots \\ Y_{ik}^{(2)} \end{pmatrix}, W=\begin{pmatrix} 1^4/n^4 & 0 & \cdots & 0 \\ 0 & 2^4/n^4 & \cdots & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \cdots & k^4/n^4 \end{pmatrix} \tag{6}$$

Then compute $\widehat{\beta}_i = (D^T W D)^{-1} D^T W Y_i^{(2)}$. Finally, the second derivative estimator of $r(x_i)$ is obtained as $\widehat{r}^{(2)}_{LowLSR}(x_i) = e_1^T \widehat{\beta}_i / l^2$, where $e_1 = (1 \quad 0 \quad 0)^T$, and l is the length of the estimation interval.

The authors developed the R package *npregderiv* [7] that includes the function *reg_2derivWL*, which allows estimating the second derivative of a regression function in the interior region by the method of Wang and Lin [6]. This function may be helpful for other researchers who work on similar problems.

The LowLSR algorithm outlined above for the interior region can be extended for the boundary regions, that is for $2 \le i \le k$ (left boundary) and $n-k+1 \le i \le n-1$ (right boundary), after replacing k in (6) by, respectively, $i-1$ and $n-i$. However, the method's accuracy of estimation in the boundary regions decreases as one approaches the boundary endpoints because the number of the data points involved in the computation decreases. The inferior quality of estimation in the boundary region is an acknowledged problem for the other methods, not only for LowLSR.

Simulation Example. In this example, the authors compare the performances of the LowLSR and ordinary polynomial LSR methods in estimating the second derivative of the function $r(x) = 32(1-2x)e^{-8(1-2x)^2}, x \in [0,1]$. The same function appears in [6] and [7]. A data set of size $n = 500$ is generated using the design points $x_i = i/n$, $i = 1, \cdots, n$, and the error terms $\varepsilon_i \sim N(0, 0.2^2)$. The authors selected $k = 0.1n = 50$.

Fig. 2(a) shows the regression function and the generated data points. The authors ended up using the ninth degree polynomial LSR method that produced the fit with all regression coefficients being significant and the adjusted R^2, the frequently used measure of goodness of fit, being equal to 0.9916, which indicates the almost perfect fit. Using the lower orders of the polynomial resulted in the poorer fits with the significantly lower values of the adjusted R^2. The fitted curve is not shown in Fig. 2(a) to not overload the graph, but it is quite close to the true function.

Fig. 2(b) shows the true second derivative function along with the LowLSR and LSR estimates in the interior region and, partially, in the boundaries. The LSR estimate is obtained after taking the second derivative of the original ninth-degree regression fit. In the interior region, the LowLSR fit closely tracks the true curve, whereas the LSR fit is not that close to it. For numerical assessment of the performances of the LowLSR and LSR methods in the interior, the following measure of error is introduced:

$$E = \frac{1}{n-2k}\sum_{i=k+1}^{n-k}\left|\widehat{r}^{(2)}(x_i) - r^{(2)}(x_i)\right|,$$

where $r^{(2)}(x_i)$ and $\widehat{r}^{(2)}(x_i)$ are the true second derivative function and an estimate, correspondingly, evaluated at x_i. The errors for LowLSR and LSR, respectively, are 31.96 and 100.29. The LowLSR fit is reasonably close to the function's second derivative in the boundary regions only for the values of x within, about, 0.05 from the left and right boundary borders. The differentiated LSR estimate is completely off in both boundary regions.

To conclude, in this example the authors found that the LowLSR method produces a quite accurate second derivative estimate of the true second derivative of the considered function in the interior region, whereas the method's performance gets worse in the boundary. The twice-differentiated ninth-degree polynomial LSR fit is inaccurate in both the interior and boundary areas.

3. Results

In this section, a detailed description of the approach in estimating the thin film residual stress using the nonparametric methods is provided. The steps of the estimation algorithm are summarized in the last subsection of this section for convenience.

For illustration purposes, the analysis results for only one substrate used in the experiment (wafer 40) are included. The results for the other substrates are found to be similar. The data frame *wafer40* in the R package *npregderiv* [7] contains the values of the design points x_i, $i = 1, \ldots, n = 7755$, and the corresponding deflection measurements for wafer 40 in the cases of the 0° and 90° scan orientations.

3.1. The scatterplots of the deflection measurements

Fig. 3 shows two scatter plots in one panel corresponding to the deflection measurements taken at the design points before and after the film deposition in the direction of 0°. In both cases, the originally measured deflection values were shifted in the vertical direction so that the deflections at $x_1 = 0$ are equal to zero. This vertical shift allows comparing the extent of deformation before and after the SiC film deposition and does not affect estimating the corresponding second derivative functions. In each case, the data points appear to line up along the smooth curves with trivial scatter, which facilitates estimating the second derivatives of the corresponding deflection functions.

For comparing the performances of LSR and LowLSR for the data, the authors first followed the traditionally used steps in analyzing the deflection measurements. That is, the polynomial LSR estimates of the deflection functions were found by, preferably, using the polynomials of the second order. The obtained fits were then differentiated twice for obtaining the corresponding curvature estimates.

For the data on 0° before the film deposition, the following quadratic regression fit was obtained:

$$\widehat{y}_1(x) = 0.1708 + 0.3911x - 0.01854x^2, \tag{7}$$

where x is the position (in mm) and $\widehat{y}_1$ is the deflection estimate (in μm). All coefficients in the above equations are statistically significant. The adjusted R^2 is equal to 0.9999.

On the graph, the quadratic fit (7) is not seen since it is masked by the trend formed by the corresponding data points. The median value of the relative absolute deviation computed as $median\left|\frac{\widehat{y}_1(x_i)-y_i}{y_i}\right|, i = 1, \cdots, n,$ constitutes 1.08%.

For the data on 0° after the film deposition, the fitted quadratic curve did not follow the corresponding points well, so the authors used the cubic regression method that produced the following result:

$$\widehat{y}_2(x) = 1.299 + 2.740x + 0.05116x^2 - 0.002157x^3, \tag{8}$$

where $\widehat{y}_2$ is the corresponding deflection estimate (in μm). All the coefficients are significant, and the value of the adjusted R^2 is 0.9995. The median relative absolute deviation, in this case, is 0.86%.

Thus, the weakness of using the second-order LSR for estimating a deflection function by default is that the fitted quadratic might not capture the trend in the deflection measurements properly. A subsequent increase of the polynomial order might be required. This makes the procedure subjective, time-consuming, and difficult to automate.

Fig. 4 shows the same type of information as Fig. 3 in the case of the 90° scan direction. The polynomial LSR fits are not computed in the 90° case for conciseness. In fact, in what follows, the superiority of using LowLSR over LSR in the pointwise curvature assessment is demonstrated.

All graphs in Figs. 3 and 4 show the concave functions. This indicates the compressive stress in the SiC film along the 0° and 90° scan

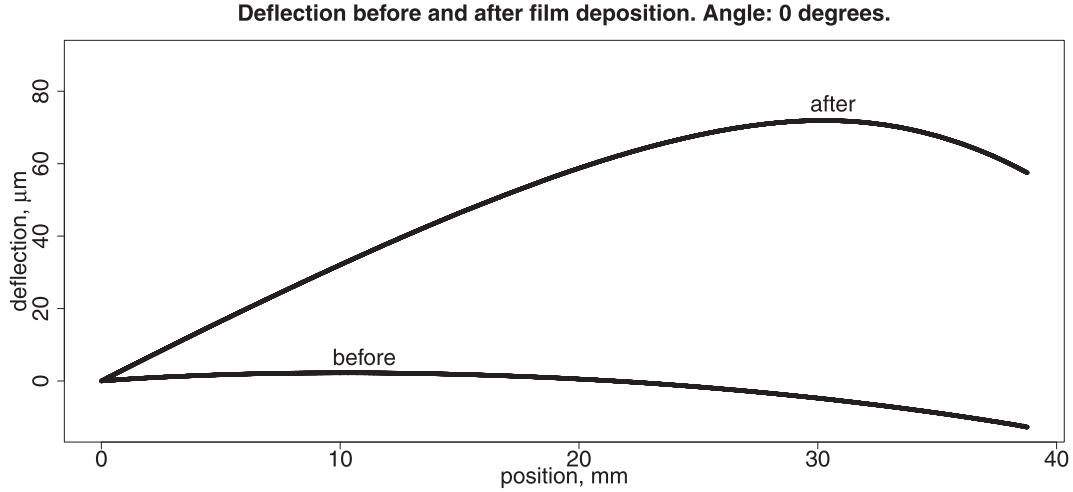


Fig. 3. Substrate deflection before and after the SiC film deposition at 0°.

directions.

3.2. Estimating the second derivatives of the deflection functions

In each of the considered settings, the authors implemented the LowLSR method based on k that is equal to the rounded value of $0.1n$ for all i in the interior region. Such a choice is within the range of the k values used in [6]. More specifically, in this case, $k = 776$. Thus, the interior region consists of the points with the values of x_i (4) for $777 \leq i \leq 6979$. Because of the boundary estimation problems discussed in Section 2.2.2, the authors only computed the LowLSR estimates in the interior region. It is worth mentioning that the data from the boundary intervals are still used for computing the second derivative estimate of $w(x)$ in the interior.

In Fig. 5, the estimates of the second derivative of the deflection function in the case of 0° (a) before and (b) after the film deposition are plotted. The solid curves show $\widehat{w}^{(2)}_{LowLSR}$, the LowLSR estimate, whereas the dashed lines correspond to the LSR estimate obtained after differentiating the LSR fits (7) and (8) correspondingly to each of the shown cases.

In Fig. 5(a), which corresponds to the case of 0° before the film deposition, the LowLSR second derivative estimate scatter around the constant LSR estimate of -0.03708 m^{-1}. The maximum relative absolute deviation of the LowLSR estimate from the constant fit is 25%. Overall, it appears from Fig. 5(a) that the assumption of the constant curvature is not that unrealistic in the considered setting. However, it is anticipated that the LowLSR estimate provides a more accurate pointwise assessment of the substrate curvature along the x-direction.

It is apparent from Fig. 5(b), which corresponds to the case of 0° after the film deposition, that the assumption of the constant curvature is unreasonable in that setting. The linear curvature estimate, obtained after differentiating the cubic regression estimate (8), does not match the LowLSR estimate. The two estimates are especially different at the edges. Thus, even though the cubic fit seems to be a good estimate of the deflection function on its own, the second derivative of that fit does not appear to be an accurate estimate of the curvature.

Fig. 6 shows in one panel both second derivative LowLSR estimates of the deflection functions in the 0° case before and after the SiC film deposition. The difference between the two estimates is also shown on the graph as in equation (2). In the united scale, it is evident that at 0° the second derivative of the deflection function before the film deposition is fairly constant along the x-axis and relatively small in magnitude. After the film deposition, the curvature appears to be mostly constant and close to zero for $x < 10$ mm and then gradually falls to the value of about -0.4 m^{-1} at the wafer flat. This result follows from Fig. 3 (b) where the deflection function is about linear for $x < 10$ mm.

Fig. 7 shows the LowLSR second derivative estimates in the case of the 90° scan direction (a) before and (b) after the thin film deposition. There is an apparent drifting downtrend in the curve in Fig. 7(a). Indeed, the curvature slopes down from about -0.037 m^{-1} at $x = 5$ mm to -0.042 m^{-1} at $x = 35$ mm. In the after deposition case, the curvature appears to be somewhat larger (in absolute value) in the middle of the substrate and is more or less constant in the other regions with an average value of about -0.23 m^{-1}.

Fig. 8 shows on one graph the LowLSR second derivative estimates of deflection before and after the film deposition and their difference in the case of the 90° scan direction. The curvature change averages to about -0.19 m^{-1}. Nevertheless, the LowLSR curvature change pointwise estimate is anticipated to be more accurate compared to this average value.

3.3. Estimating the thickness profiles

Fig. 9 displays the data points and the local linear estimates (LLE, see [15]) of the SiC film thickness functions $t_f(x)$ in the (a) 0° and (b) 90° scan directions. The LLE are computed based on the bandwidth of 2 mm using the function *loclin* from the R package *OSCV* [19]. In the 0° case, the SiC film thickness increases almost linearly in the direction of the CVD reactants flow. This is consistent with the corresponding result in [20]. In the 90° direction, the thickness function appears to be concave and fairly symmetrical with the peak located around $x = 20$ mm. In fact, the obtained curve shapes are consistent with the prior knowledge of the specific CVD reactor (see [21]).

The film thickness varies by 75% and 37% along the 0° and 90° scan directions, respectively. Such a great variation is unusual in practice for the production film deposition, and can be possibly avoided with wafer rotation during deposition.

3.4. Computing the residual stress estimates

Ultimately, Stoney's equation (1) was used to estimate the resulting SiC film residual stress as a function of the x-coordinate in the 0° and 90° scan directions. In each case, the estimated film thickness (see Fig. 9) and the curvature change (see Figs. 6 and 8) were plugged into equation (1). The other components of the residual stress computation are the substrate thickness $t_s = 275$ μm, Young's modulus $E_s = 130$ GPa, and Poisson's ratio $\nu_s = 0.279$. The resulting estimates of the residual stress in the 0° and 90° directions are shown in Fig. 10 (a) and (b), respectively.

At 0°, the residual stress linearly drops from, essentially, 0 MPa to

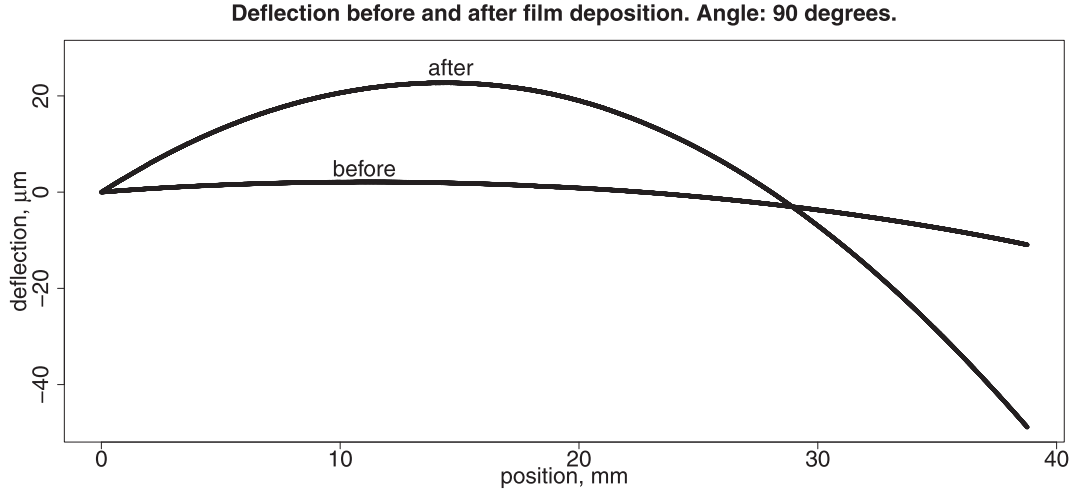
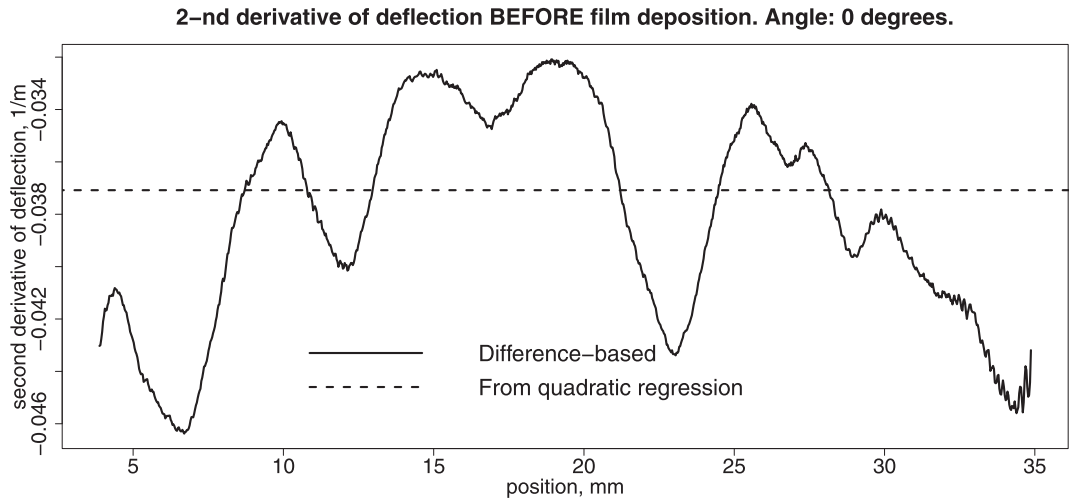


Fig. 4. Substrate deflection before and after the SiC film deposition at 90°.

(a)



(b)

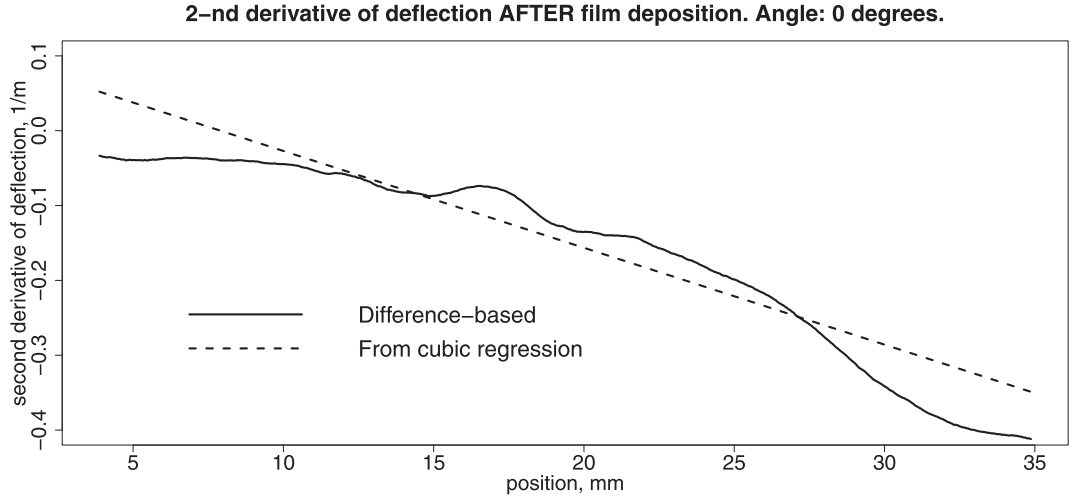


Fig. 5. Estimated second derivatives of the deflection functions (a) before and (b) after the SiC film deposition at 0°.

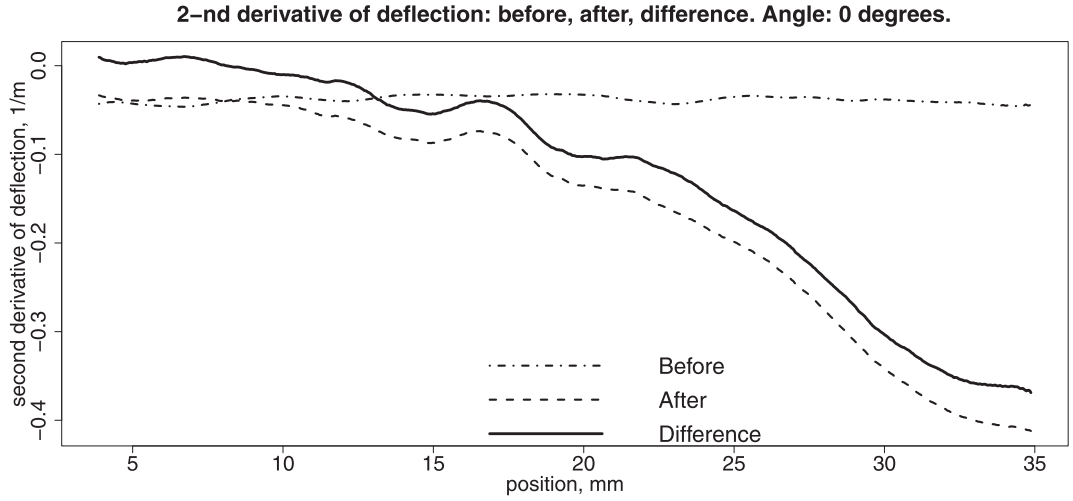
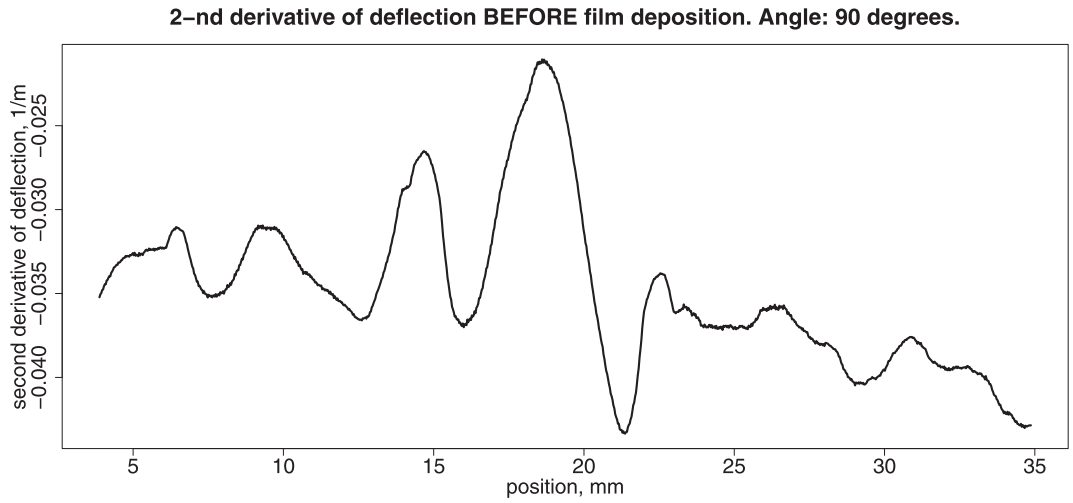


Fig. 6. Estimated second derivatives of the deflection functions before and after the SiC film deposition and their difference in the 0° scan direction.

(a)



(b)

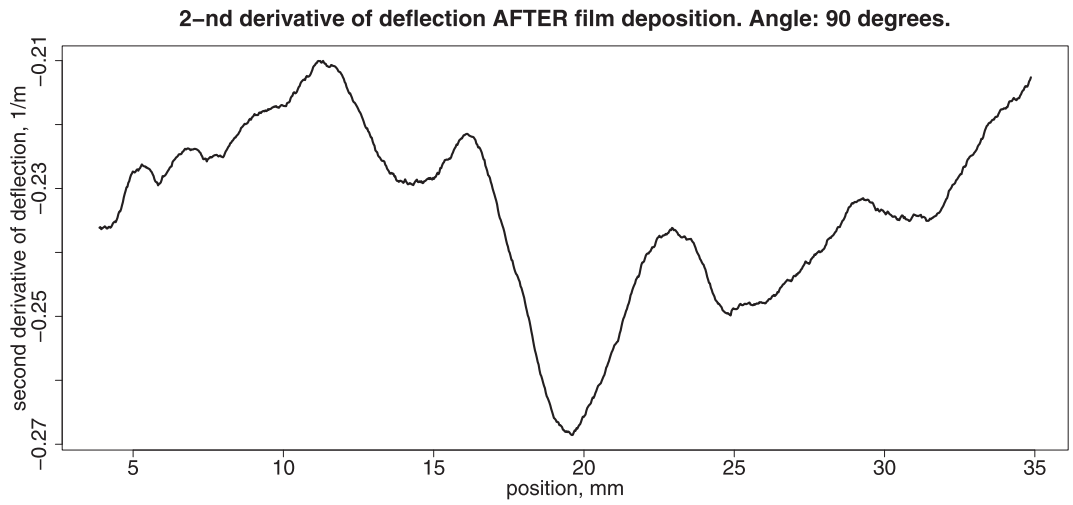


Fig. 7. Estimated second derivatives of the deflection functions (a) before and (b) after the SiC film deposition at 90°.

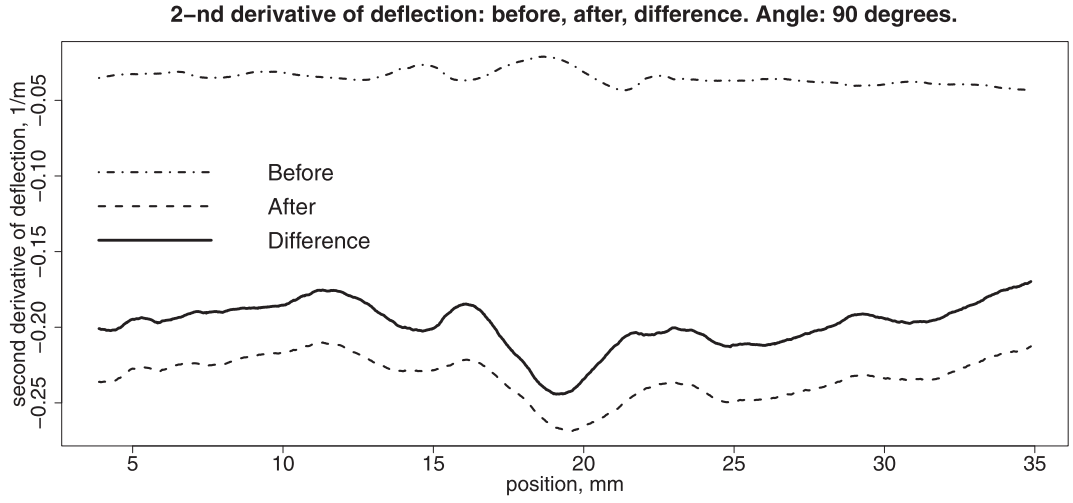


Fig. 8. Estimated second derivatives of the deflection functions before and after the SiC film deposition and their difference in the 90° scan direction.

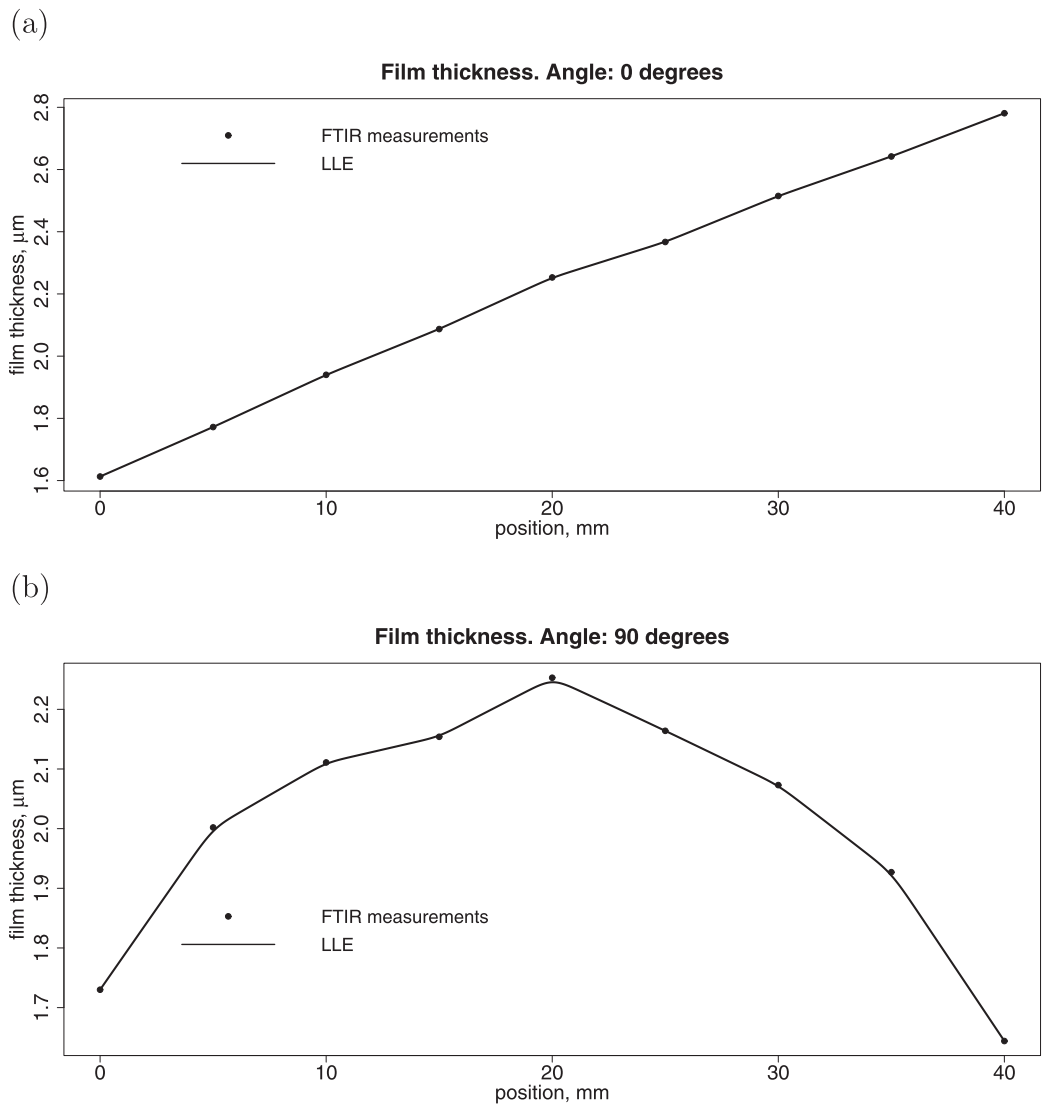


Fig. 9. Film thickness in the (a) 0° and (b) 90° scan directions.

about −300 MPa. Almost zero stress around $x = 0$ mm is consistent with the obtained almost trivial value of the estimate of $\Delta\kappa(0)$ (see Fig. 6). The residual stress increases (in absolute value) with the SiC film thickness. In the 90° case, the residual stress changes from about −250 MPa to −190 MPa, and is the greatest (in magnitude) in the middle area of the wafer.

It is worth mentioning that the values of the residual stress at the center of the wafer computed in the 0° and 90° directions do not match, since the residual stress in the film is not equibiaxial. The obtained mismatch supports the rationality of using Stoney's equation (1) locally.

3.5. Algorithm for nonparametric estimation of the residual stress (summary)

For convenience, the steps of the algorithm for estimating the residual stress in thin films are summarized below.

Given the set of n wafer deflection measurements taken by a surface profilometer at the evenly spaced design points along the x-axis that passes through the wafer center in the selected (radial) direction (see Fig. 1(b)), estimate the corresponding residual stress function in the interior region by following the steps outlined below.

Step 1: Defining the interior region. Compute $k = round(0.1n)$. Then the interior region is defined as the set of the data points with the x-values in the range $x_{k+1} \le x \le x_{n-k}$.

Step 2: Obtaining the estimate of the curvature change function. Use the R function *reg_2derivWL* from the package *nprederiv* to compute two curvature function estimates at the design points within the interior region based on the deflection measurement data before and after the film deposition. Then compute the curvature change estimate as the difference of the two functions, as in (2).

Step 3: Estimating the film thickness function. From the film thickness measurements taken by the FTIR spectrometer in the x-direction, use the LLE to estimate the thickness function in the interior region. One may use the R function *loclin* from the R package *OSCV*.

Step 4: Computing the estimate of the residual stress. Obtain the ultimate estimate of the residual stress profile in the interior region by plugging the curvature change estimate and the thickness function estimate obtained in Steps 2 and 3, respectively, into Stoney's equation (1).

The above algorithm can be extended for the x-values in the left and/or right boundary regions. See Section 2.2.2 for the details of estimating the curvature in the boundary region. Also, observe that the outlined steps can be used for estimating the residual stress in thin films not only in the radial direction but also along any line of the substrate surface.

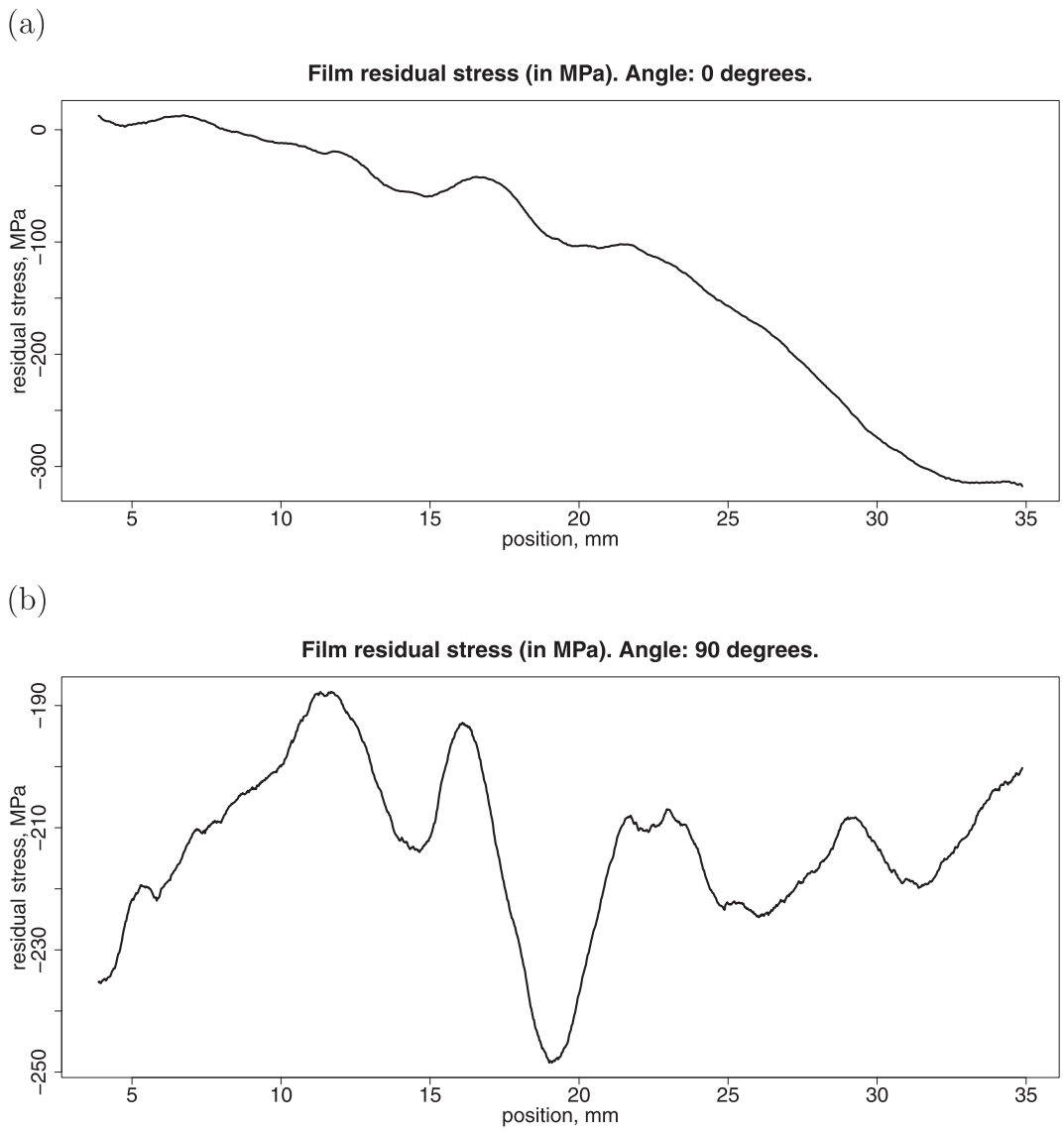


Fig. 10. Estimated SiC film residual stress along the x-coordinate for the (a) 0° and (b) 90° scan directions.

4. Discussion

For future research, it appears to be more relevant to plug into the Stoney's equation at each point (x,y) on the substrate the largest in the magnitude of two principal curvatures $\kappa_1(x,y)$ and $\kappa_2(x,y)$ of $w(x,y) = w_2(x,y) - w_1(x,y)$, where $w_1(x,y)$ and $w_2(x,y)$ denote the deflection functions before and after the film deposition, respectively. These principal curvatures, in the case when the first derivatives of w are assumed to be 0, are computed as the eigenvalues of the Hessian matrix $H(x,y) = \begin{pmatrix} w_{xx} & w_{xy} \\ w_{yx} & w_{yy} \end{pmatrix}$ of the second partial derivatives of w at the point (x,y) (see [22]). The principal curvatures $\kappa_1(x,y)$ and $\kappa_2(x,y)$ represent the maximal and minimal directional curvatures of w at the point (x,y), whereas the corresponding eigenvectors specify the directions where these curvatures are achieved. Since these eigenvectors are known to be orthogonal, the eigenvector corresponding to the smaller in magnitude principal curvature represents the most likely direction of the film fracture at the considered point. In a more general case when the first derivatives of $w(x,y)$ cannot be taken equal to zero, the principal curvatures are computed as the eigenvalues of the shape operator S of the surface parameterized by $(x,y,w(x,y))$. The reader is referred to [22] for details. Implementing the above ideas requires adapting the existing or developing the new methods of estimating the mixed second derivative of the difference of the deflection functions before and after the film deposition and elaborating the experimental designs that allow for taking all necessary measurements.

In addition to the contact profilometer, a developed methodology of estimating the thin film residual stress can also be applied to the wafer deflection data obtained by an optical profilometer [23].

5. Conclusions

The article outlines an algorithm for estimating the SiC film residual stress profiles in the radial directions for a round substrate in the case of the non-uniform film thickness. The algorithm is based on Stoney's equation used locally by plugging into the formula the estimated curvature change and thickness profiles along the selected radial directions. The curvature change profile is estimated by the LowLSR nonparametric statistical method developed by Wang and Lin [6]. The thickness profile is estimated using the LLE.

In this research, the LowLSR nonparametric method of the second derivative estimation was preferred because it is shown in the simulation study in [6] to outperform some of its frequently used competitors. It is also shown in the Appendix that LowLSR is more efficient than

LowLAD [14] in the case of the normally distributed error terms.

Estimating the curvature based on LowLSR requires selecting k, the method's smoothing parameter. The authors of the method [6] derived an expression for an asymptotically optimal value of k that is currently of no practical value. The authors empirically found that $k = 0.1n$ works well in a variety of settings. Such a choice is within the range of the k values used in [6]. Developing a more sophisticated data-driven rule for selecting k that takes into account its $n^{12/13}$ rate (see (5)) would have been beneficial. A sort of plug-in rule similar to [24] can be elaborated.

The benefit of estimating the curvature via LowLSR instead of the traditionally used approach of first estimating the deflection function based on the polynomial LSR method and then taking the second derivative of the obtained fit is illustrated in the simulation example included in Section 2.2.2. Thus, LowLSR is anticipated to provide a more accurate pointwise curvature change assessment in a selected radial direction compared to the estimate obtained after differentiating the polynomial LSR fit.

This study illustrated assessing the thin film residual stress only in the 0° and 90° radial directions (see Fig. 1(b)). Any other direction can be treated in the same way, if needed.

The estimated film residual stress profiles differ substantially for 0° and 90° scan orientations. At 0° the stress drops almost linearly towards the opposite side of the wafer from about 0 MPa to −300 MPa (see Fig. 10(a)) That is, the film residual stress in the 0° case is found to be the greatest (in magnitude) around the wafer flat where the film is the thickest. At 90°, the residual stress changes in the range (−250, −190) MPa and is the greatest (in magnitude) in the middle of the wafer (see Fig. 10(b)). The mismatch of the residual stress values computed in the 0° and 90° directions at the center of the wafer validates using Stoney's equation locally for an adequate assessment of the residual stress at a given point and in a selected direction.

For a more accurate pointwise residual stress estimation across the substrate, one should be able to find the principal curvatures at any selected point in the corresponding principal directions. This requires adapting the existing or developing the new algorithms of estimating the second mixed derivative of a deflection function and elaborating the experimental designs that allow for all corresponding measurements. These might be the topics of future research efforts.

All computations in this research are performed in the R package, which is a commonly used programming language in statistics and applications (see [25]). In particular, the R package *npregderiv* [7] contains the data set *wafer40* that was used for illustration purposes in this article and the function *reg_2derivWL* that implements LowLSR [6]. The package also includes the function *reg_1derivWL* that can be used for estimating the function's first derivative by the method [6]. These functions can be useful for other researchers and/or data analysts working on similar or related problems.

CRediT authorship contribution statement

Olga Savchuk: Writing - original draft, Formal analysis, Methodology, Software, Visualization. **Alex A. Volinsky:** Writing - review & editing, Data curation, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to acknowledge Professor Saddow's group from USF for SiC films deposition, Jose Carballo for performing profilometer measurements, and Grygoriy Kravchenko for FEM data analysis. AV acknowledges support from the National Science Foundation (IRES 1358088) and the Government of the Russian Federation (2020-220-08-6662).

Appendix

In the case of the second derivative estimation of $r(x_i)$ in the regression model (3), the LowLSR [6] and LowLAD [14] methods have identical main terms in their asymptotic bias expressions. The ratio of the main terms of the asymptotic variance expressions of the LowLSR (top) and LowLAD (bottom) methods, appears to be equal to

$$R_{Var} = 2h(0)^2\sigma^2, \tag{9}$$

where $h(0) = \int_{-\infty}^{\infty} f^2(x)dx$. The ratio of the corresponding minimum asymptotic mean squared errors (AMSE) is equal to

$$R_{AMSE} = \left(2h(0)^2\sigma^2\right)^{8/13}. \tag{10}$$

It is commonly assumed that f is $N(0, \sigma^2)$. It can be shown that in this case $h(0) = \frac{1}{2\sigma\sqrt{\pi}}$. It then follows from (9) and (10), correspondingly, that $R_{Var} = \frac{1}{2\pi} \approx 0.1592$ and $R_{AMSE} = \left(\frac{1}{2\pi}\right)^{8/13} \approx 0.3227$. These results imply that the LowLSR method is more efficient than LowLAD in the case when the errors in the regression model (3) are normally distributed, which is the most frequently made assumption.

References

[1] G.G. Stoney, The tension of metallic films deposited by electrolysis, Proc. R. Soc. London Ser. A 82 (1906) 172–175.
[2] D. Ngo, X. Feng, Y. Huang, Thin film/substrate systems featuring arbitrary film thickness and misfit strain distributions. Part I: analysis for obtaining film stress from non-local curvature information, Int. J. Solids Struct. 44 (6) (2007) 1745–1754.
[3] M. Brown, Measuring Stress in Thin Film-substrate Systems Featuring Spatial Nonuniformities of Film Thickness and/or Misfit Strain. California Institute of Technology, Ph.D. Thesis, 2007.
[4] M. Zecchino, T. Cunningham, Thin film stress measurement using Dektak surface profilers, Bruker Application Note, 2010, https://www.bruker.com/fileadmin/user_upload/8-PDF-Docs/SurfaceAnalysis/StylusProfilometry/ApplicationNotes/AN516-Thin_Film_Stress_Measurement_Using_Dektak_Stylus_P.pdf.
[5] H.-G. Müller, U. Standtmüller, T. Schmitt, Bandwidth choice and confidence intervals for derivatives of noisy data, Biometrika 74 (4) (1987) 743–749.
[6] W. Wang, L. Lin, Derivative estimation based on difference sequence via locally weighted least squares regression, J. Mach. Learn. Res. 16 (2015) 2617–2641.
[7] O. Savchuk, A. Volinsky, npregderiv: Nonparametric estimation of the derivatives of the regression function, R CRAN package version 1.0, 2020, https://cran.r-project.org/web/packages/npregderiv/index.html.
[8] S. Zhou, D.A. Wolfe, On derivative estimation in spline regression, Statistica Sinica 10 (1) (2000) 93–108.
[9] D. Ruppert, M.P. Wand, Multivariate locally weighted least squares regression, Ann. Stat. 22 (3) (1994) 1346–1370.

[10] J. Fan, I. Gijbels, Local Polynomial Modelling and its Applications, Chapmen & Hall, London, 1996.
[11] K. De Brabanter, J. De Brabanter, B. De Moor, I. Gijbels, Derivative estimation with local polynomial fitting, J. Mach. Learn. Res. 14 (1) (2013) 281–301.
[12] Y. Liu, K. De Brabanter, Derivative estimation in random design, in: 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montreal, Canada, 2018.
[13] W. Dai, T. Tong, M.G. Genton, Optimal estimation of derivatives in nonparametric regression, J. Mach. Learn. Res. 17 (2016) 1–25.
[14] W. Wang, P. Yu, L. Lin, T. Tong, Robust estimation of derivatives using locally weighted least absolute deviation regression, J. Mach. Learn. Res. 20 (2019) 1–49.
[15] W.S. Cleveland, Robust locally weighted regression and smoothing scatterplots, J. Am. Stat. Assoc. 74 (368) (1979) 829–836.
[16] F. Jiang, S. Chen, Y. Leng, N. Huang, Effect of wafer size on the film internal stress measurement by wafer curvature method, J. Wuhan Univ. Technol.-Mater. Sci. Ed. 31 (2016) 93–99.
[17] FLX 2320-S Thermal Stress Measurement, KLA-Tencor Application Note, 2019, https://tohotechnology.com/wp-content/uploads/2019/09/Toho_FLX2320S_D1.pdf.
[18] J. Carballo, Residual Stress Analysis in 3C-SiC Thin Films by Substrate, Curvature Method, University of South Florida M.S. Thesis, 2010, https://scholarcommons.usf.edu/etd/1590/..
[19] O. Savchuk, OSCV: One-Sided Cross-Validation, R CRAN package version 1.0, 2017, https://cran.r-project.org/web/packages/OSCV/index.html.
[20] A. Volinsky, G. Kravchenko, P. Waters, J. Reddy, C. Locke, C. Frewin, S. Saddow, Residual stress in CVD-grown 3C-SiC films on Si substrates, MRS Proc. 1069 (2008) 1069-D03-05.
[21] M. Smith, Design and development of a silicon carbide chemical vapor deposition reactor, University of South Florida M.S. Thesis, 2003, https://digital.lib.usf.edu/SFS0024841/00001.
[22] M. Do Carmo, Differential Geometry of Curves and Surfaces, revised and updated second edition, Dover Publications, Inc., 2016, p. 510.
[23] A. Besnard, M.R. Ardigo, L. Imhoff, P. Jacquet, Curvature radius measurement by optical profiler and determination of the residual stress in thin films, Appl. Surf. Sci. 487 (2019) 356–361.
[24] D. Ruppert, S. Sheather, M. Wand, An effective bandwidth selector for local least squares regression, J. Am. Stat. Assoc. 90 (432) (1995) 1257–1270.
[25] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2013 http://www.R-project.org/.