1. Introduction

Nowadays, mechanical designers and engineers of elastic structures at ultra-small scales face an interesting challenge; traditional flexures with several components, mechanical joints/welds and linkages are almost impossible to manufacture with existing microsystem fabrication technologies. Therefore, the majority of mechanical components, elements and building blocks are based on micro/nano machined elastic flexures [1]. Despite the complexity of modeling and analyzing these systems, their design uses the averaged relation between stress and strain, which requires a relatively accurate knowledge of their effective elastic properties, specifically the effective elastic modulus. Experimental results show that the elastic properties are constant at length scales of meters down to micrometers [2]. However, in order to increase the performance, i.e. sensitivity and dynamic range, the dimensions of mechanical devices have been scaled down towards a few nanometers. Consequently, high performances such as single-electron tunneling [3], sub-attonewton force sensing [4], and sub-femtometer displacement sensing [5] have been successfully achieved. Unlike at micron and higher length scales, at sub-micron and nanometer length scales the effective elastic behavior shows strong scale-dependent behavior, meaning that the elastic properties are no longer constant, but a function of length scale. For all the application examples above, the high performance was achieved due to the mechanical response of a nanosystem, which strongly depends on the effective elastic properties. Therefore, a clear understanding of the scale-dependent behavior is important for the design and performance of nanosystems.

1.1. Scale-dependence

The existence of scale-dependent behavior has been confirmed by experimental measurements, including resonance frequency tests [6], tensile testing in scanning electron microscope
(SEM) [7], transmission electron microscope (TEM) [8, 9], atomic force microscope (AFM) [10, 11] and nanoindenter [12] and also theoretical investigations, including \textit{ab initio} and density functional theory (DFT) [13, 14, 15], molecular dynamics (MD) [16, 17, 18, 19] and modifications to continuum theory [20, 21, 22, 23]. Although scale-dependence has been observed by both theory and experiment, a considerable discrepancy still remains between the experiments and models. As an example, in mono-crystalline [110] silicon nanowires and nanocantilevers the experimental scale-dependence has been observed at about 150 nm [24, 11, 8, 6, 25], whereas theoretical studies do not agree with any scale-dependence larger than 10 nm. In order to highlight the inescapable discrepancy between experiment and theory, the results of recent experimentally measured effective Young’s modulus \( E \) and those extracted from theories for silicon nanocantilevers and nanowires were collected from relevant literature and are illustrated in Fig. 1. The figure demonstrates that a large gap exists between the theoretical predictions and experimental observations. The smallest silicon nanowire with a diameter of about 12 nm that has been experimentally tested by Han et al. [26] shows an effective elastic modulus which is about one third of the bulk value, while the largest theoretically modeled (mainly atomistic simulations) silicon nanowire (diameter \( \sim 10 \) nm) shows an effective elastic modulus that is closer to the bulk value [26]. It can also be observed from Fig. 1, that there is lack of data, both in experiment and theory: 1) there is no computational data for structures in the range from 10 to 150 nm, so no direct comparison between experiments and theory at about 100 nm can be performed. The reason is that it is computationally extremely expensive and impossible to model the atomistic systems as large as the ones experimentally tested. Although quasi-continuum approaches have been developed as computationally efficient methods [27, 28], they are lacking important phenomena, such as surface reconstruction, defects emitted from surfaces and surface stress induced phase transformations. Thus, one of the biggest challenges in the field of nanomechanics is the development of a multi-scale modeling framework, capable of performing simulations at various length scales. 2) There is no experimentally measured data below 12 nm, therefore for structures smaller than 10 nm a direct comparison between theory and experiment cannot be performed either. The main reason is that it is extremely challenging and difficult to perform nanoscale experiments with sufficient accuracy and resolution, while introducing minimum disturbances to the measured nanostructure [24]. Therefore, experimental capabilities need to be improved, enabling measurements of structures with a size of less than 10 nm.

Moreover, the observed discrepancies are not only between the simulations and experiments but also between various experimental measurements. The observed discrepancies [25, 29, 24] could be explained by: some exterior factors such as ill-defined boundary conditions and uncertainties in experiments [30], loading modes (extensional vs. bending) [31, 23], calibration and required input accuracy (i.e. variations in geometry) limitations [11] as well as intrinsic nature of inaccuracy in the experimental methods (i.e. mass effects in resonance based tests) [32, 33, 34] and the environmental effects such as surface contaminations, measurement induced-errors and native oxide layers could all be influencing the measurement at sub-micron scales, inducing a large disagreement between measurement data. On the other hand, there are intrinsic effects which become significant at sub-micron scales such as surface stress, surface elasticity, nonlinear bulk effects, surface reconstructions, crystal defects and fabrication induced imperfections that make the scale-dependent phenomena more complicated.
Due to their small sizes, nanosystems predominantly have high surface-to-volume-ratio, thus, unlike micron size systems, the influence of surface stress cannot be neglected. Therefore, the scale-dependence of elastic behavior is generally attributed to surface effects, including surface elasticity [35, 20, 23] and surface stress [36, 37], surface oxidation [31, 29, 38] and surface contamination [39]. Besides the intrinsic surface effects, it was speculated that surface geometry and cross section are important factors. Most simulations have been performed on nanowires with rectangular cross section, which is significantly simplified compared to actual cross sections that are closer to rhombic or pentagonal configurations [40]. In order to examine to what extent these effects are significant, McDowell et al. [31] using atomistic simulations, studied the scale-dependent elastic behavior of nanowires with various cross sections and with various surface steps. The conclusion was that, although these factors influence \( E \) to some extent, they alone are insufficient to explain the experimentally observed trends.

In addition to surface geometry, the loading method can also influence \( E \). The extracted \( E \) in the extensional mode is different from the one extracted in the bending mode [16]. Consequently, there are two types of experimental characterizations: those measuring \( E \) in the extensional mode, such as uniaxial tensile loading tests [41, 42, 9], and those in the bending mode such as resonance frequency based [6] tests or bending tests inside an atomic force microscopes [10, 11]. Most of the research today focuses on the bending mode of \( E \) because it is more sensitive to surface stress and surface elasticity effects and because of its importance in sensing and actuating applications. Again the question arises: at which scale does the difference between bending and extensional modes become significant? To answer this question, McDowell et al. [16] investigated the effect theoretically and found out that it is influential only for dimensions below approximately 8 nm. This indicates that other factors beyond the above effects play a role in experimentally observed nanosystem scale-dependent elastic behavior. Therefore, the question remains: why does \( E \) start to decrease already at sizes between 10 nm and 150 nm?

A couple of issues have been raised by researchers in order to reduce the gap between the experiments and the theory [24, 11, 39, 16, 25].

A recently recognized substantial effect, which has not yet been considered, is accounting for the fact that experimentally measured nanosystems are not perfect (such as existence of surface defects, native oxide layers, contaminations etc.) and not defect free (such as single crystal defects, fabrication induced imperfections etc.) in contrast to the perfect mono-crystalline structures studied using atomistic simulations [19, 17, 14, 13]. The goal of this chapter is to try to determine to what extent these effects influence the effective elastic modulus of nanosystems, to try to reduce the existing gap between the experiments and theory of scale-dependence, and to determine the origin of the scale-dependence in silicon nanosystems.

1.2. Outline of the chapter

This chapter is organized as follows: Section 2 describes the fabrication process for micro/nano cantilevers, used for experimental investigations of scale-dependent \( E \), including the inspection of the fabricated structures. In Section 3 the experimental determination of \( E \) is presented and discussed. Section 4 is focused on the theoretical investigations of surface effects and native
Figure 1. (a) Illustration of the scale-dependence of mono-crystalline silicon effective elastic modulus $E$ obtained via computations and experimental observations [29]. The inset shows results for the scale less than 10 nm, indicating that all simulations are limited to scales of less than 10 nm due to complexity and expense of computations. (b) A logarithmic plot of (a) to further illustrate the differences between the simulation and the experimental results. Both show sharp drops near the down-scaling side; the question is why discrepancies exist? (Reprinted with permission from [29]. Copyright 2010 IOP Publishing Ltd.)
oxide layers on the scale-dependent $\hat{E}$, including molecular dynamics calculations as well as semi-continuum approaches, to show to what extent these effects are responsible for the observed scale-dependence. Section 5 discusses the possible effects of defects that may contribute to the reduction of $\hat{E}$ at larger scales. In order to show the extent that defects play a role in scale dependence, molecular dynamics is utilized to quantify the effects of defects on $\hat{E}$. Finally in Section 6, a summary, conclusions and recommendations for future work on scale-dependence of $\hat{E}$ are presented.

2. Fabrication process and inspection of mono-crystalline silicon micro/nanosystems

Micro and nanocantilevers, double clamped structures and plates are widely fabricated with top-down approaches [6, 32], while nanowires and nanotubes are produced invoking bottom-up techniques [43]. This section describes the fabrication process and inspection of the silicon cantilevers and double clamped beams studied in this chapter. The characteristic of the acquired structures are 1) their thickness has to be accurately known, 2) their surface has to be very smooth and 3) the pre-bending due to releasing the structures, or in other words, the gap between the substrate and the structures has to be known. Inspections of the test structures were conducted in order to evaluate these criteria. The thickness of the device and substrate layers were measured by ellipsometry. AFM roughness measurement together with the results of ellipsometry were used for investigating the roughness of the structures. And finally, white light interferometry was used to measure the initial curvature of the structures.

The mono-crystalline silicon cantilevers and double clamped beams were fabricated on (100) Silicon On Insulator (SOI) wafers from Soitech, using the Smartcut® process to form the buried oxide and the silicon device layer. The used wafers are shipped with 1 μm (measured 1019 nm) thick buried oxide and a 340 nm thick low boron-doped silicon device layer. From these original wafers, thicker layers were obtained by epitaxial growth, while for the thinner ones they were thermally oxidized and then etched to the desired thicknesses. The surface roughness measured by AFM revealed a smooth surface (RMS of ~ 0.095 nm) on both the original wafer, as well as the epitaxied sample and the oxidation-etched samples (Fig. 2).

The device thicknesses were controlled by careful timing of the epitaxy and the oxidation processes. Test wafers were added to the processing batch to verify the thicknesses; the modified SOI layers were then precisely determined by ellipsometry, a pre-defined model is fitted onto the measured data to obtain the exact thicknesses. Typical fittings and their results are shown in Fig. 3. Samples with desired thickness were first spin coated with HMDS as a primer that serves as an adhesion promoter for the photoresist. The recipe is: 5000 RPM, 2 min at 200°C bake in a hot plate, then coat with photoresist AZ5214 of 1.25 μm, 5000 RPM, 2 min at 120°C and then pattern in a lithography step (Fig. 4.a).

The wafers were developed in pure MF321 for about 85 seconds and 30 seconds in water. The sample is then etched in a SF6 based plasma etcher to pattern the top silicon layer until reaching the BOX layer (Fig. 4.b). To release the cantilever, the residue of photoresist is removed in nitric
acid and the BOX is etched in Hydrofluoric (HF) solution, in order to release the structures (Fig. 4.c). Drying the structures directly after etching with HF and diluting in water can cause stiction of the structures to the substrate due to the water surface tension. In order to prevent device stiction, freeze drying or critical point drying (CPD) were used. The cantilevers were 170 to 8 μm long, 20 to 8 μm wide and 1019 to 40 nm thick. Fig. 5 shows SEM images of fabricated cantilevers and double clamped beams.

Due to the isotropic nature of the release-etching, undercuts and hidden anchors were formed at the clamping site of the cantilevers. This has an overall lengthening effect on the device. The extra length was determined with the optical microscope, as shown in Fig. 6.a. In addition, as shown in the HF etching of the BOX, the bonding surfaces of the SOI wafers were in the middle of the BOX layer about 340 nm below the top silicon oxide interface (Fig. 6.b and Fig. 6.c). This eliminates the bonding-induced defects on the top silicon and greatly improves the surface quality of the device silicon. Residual stress, due to the inconsistency of both the thermal expansion coefficient and the crystal lattice period between the substrate and the thin film, is unavoidable in surface micromachining techniques [44]. The compressive residual stress in relatively thin double clamped beams can create pre-bending. Moreover, it can create built-in moments, which in released cantilevers cause them to curl out of plane. The cantilever has one end free and therefore, it can partially release the stress. The consequence of this stress releasing is an upward or downward curvature of the cantilever.

Bending profiles of each structure before the experiment were measured using a Wyko NT3300 optical surface profiler that utilizes white light interferometry for high resolution three-dimensional (3D) surface measurements. Fig. 7.a shows the white light interferometry measurements for a cantilever which is approximately flat. One of the most important parameters that affects the accuracy of experiments is the curvature of the structures. Therefore, for each cantilever that is bent during the releasing process, a profile measurement was performed. This bending modifies the initial gap between the suspended structure and the substrate. Fig. 7.b and Fig. 7.c show the reconstructed 3D profiles of a cantilever and a double clamped beam with significant curvatures. Fig. 7.d shows the profile measurements in the length direction of the curled cantilever. The markers indicate the length of the cantilever.
Figure 3. Typical ellipsometer measurements of the SOI wafers with (a) 40 nm and (b) 1019 nm device layer thicknesses. The dots are the measured data, solid lines represent the expected values generated from the fitted thickness with different incident angles (angles are shown in the legend). The mean squared errors (MSE) of the fitting were less than 18.
3. Experimental measurement

Meaningful experimental measurements of scale-dependent elastic behavior from micro-
down to nano-scales have been shown to be one of the major challenges in the field of
experimental micro/nanomechanics [24, 45]. As a consequence, a majority of research con-
cerning the scale-dependent elastic response has focused on developing experimental methods
to study the phenomena [24, 41, 42, 8, 32]. One of the most common approaches to measure \( \hat{E} \)
is based on the resonance response of the system. \( \hat{E} \) is extracted via a simple Euler-Bernoulli
beam equation from a resonance frequency response acquisition [6, 46]. The main advantages
of this method, which make it popular and widely used, are 1) ease of use; besides measuring

---

**Figure 4.** Fabrication process of the cantilevers. (a) photoresist is patterned. (b) Using the photoresist as an etching
mask, the silicon underneath is etched with SF6 plasma. (c) Residue of the resist is stripped in nitric acid and the device
is submerged in HF solution to etch the oxide. The device is then put in a critical point drier (CPD) to release the sus-
pending cantilever structure.

**Figure 5.** SEM of (a) 1019 nm, (b) 340 nm, (c) 93 nm, (d) 57 nm, (e) 40 nm thick cantilevers. SEM of (f) paddle-shaped
cantilevers, (g) covered cantilevers and (h) double clamped beams, respectively.
the resonance frequency no additional modifications on the system are required, and 2) the method is very fast, while being applicable in various environments, including vacuum, gaseous and liquids. The method has some limitations, especially when being used at nano-scale. The experimental results include errors due to the uncertainty of the boundary conditions [30]. The non-ideal boundary conditions lead to a lower resonance frequency, which leads to lower estimates of $E$. Moreover, the resonance frequency depends on both the stiffness and the mass [39, 33]. It is therefore difficult to decouple solely by a resonance response the stiffness from the mass changes caused by surface contamination, native oxide and other adsorbed layers [33, 39, 34]. The smaller the nanosystem, the bigger this effect will be. In order to show the extent to which the effects of the extra mass due to contamination or native oxide are significant, we investigated the effects of surface contaminations on the resonant behavior of the cantilevers fabricated in Section 2. We have shown that the mass and the stiffening effects of the contaminations can cause significant shifts in the resonance frequency of the cantilevers, and, thus, the subsequent estimate of the effective stiffness.

Figure 6. (a) Optical microscope image of a cantilever with hidden anchor due to under etching of the base of the cantilever. (b) SEM of a hidden anchor. The cantilever was forcibly removed in order to see the hidden anchor. The inset shows the bonding surfaces of the SOI wafers were in the middle of the BOX layer. (c) Close-up view of bonding interface and silicon device layer.
The resonance frequency of the cantilevers was measured with an optical laser detection setup equipped with a vacuum chamber [39]. The cantilevers are in resonance due to Brownian motion caused by thermal noise. The resonance frequency and its shift were obtained from the Lorentzian fit of the frequency-bandwidth curves [39]. If the surface of the cantilever is clean without any contamination, such as adsorbed water and gas molecules, then the theoretical effective Young’s modulus of the cantilever can be calculated as [39]

\[ \hat{E} = \frac{0.0261 f_0^2 L^2 \rho}{t} \]  

where \( f_0 \) is the resonance frequency, \( L \) is the length of the cantilever, \( \rho \) is the density of the cantilever material and \( t \) is the thickness, respectively. In order to determine the variations in \( \hat{E} \) associated with the surface contamination, the resonance frequency of cantilevers with contaminations and with minimized contaminations were compared. Assuming that the contamination is distributed evenly along the cantilever, the variations in \( \hat{E} \) can be calculated using (only for small variations)

\[ \Delta \hat{E} = 2 \frac{\hat{E}}{f_0} \Delta f \]  

Figure 7. Reconstructed 3D image of (a) a cantilever which is approximately flat, (b) a curved double clamped beam and (c) a curled cantilever using white light interferometry. (d) curvature profile measurement of the cantilever in (c).
where $\Delta \bar{E}$ is the variation in $\bar{E}$, and $\Delta f$ is the variation in the resonance frequency due to the contaminations. To show experimentally the effect of ambient contamination, we analyzed the resonance frequency shift of cantilevers by the following procedure (similar to the work in [39] for inaccuracies in frequency-based mass sensors). After fabrication of the cantilevers, their resonance frequencies, $f_{bv}$, were measured in ambient air, to allow for the adsorption of different gases and water molecules and their saturations. At this stage, the masses and stiffness from surface contaminations are added to the newly fabricated cantilevers and are referred to as the added mass-stiffness effects. The cantilevers were then placed into a vacuum ($10^{-6}$ mbar) for a sufficiently long time to allow for degassing and desorption at the surfaces of the cantilever (reduced added mass-stiffness effects). Further, the resonance frequencies were measured in a vacuum ($f_v$). Then, the vacuum chamber was vented with ambient air and the resonance frequencies were measured immediately after the pressure reached atmospheric pressure ($f_{av}$). The measurements were done within a few minutes in order to ensure minimal re-adsorption, while providing the same amount of air mass loading compared to the measurements in the first step ($f_{bv}$). This procedure was repeated for 1019, 340 and 93 nm thick cantilevers. A typical resonance frequency measurement with the procedure above is illustrated in Fig. 8 for a 93 nm thick, 8 $\mu$m wide and 8 $\mu$m long cantilever. A 20.9 KHz resonance frequency shift was observed between the resonance frequency before and after vacuum. For thicker cantilevers the resonance frequency shift was less, due to their lower surface-to-volume ratio and consequently their being less sensitive. Table 1 shows the results of the resonance frequency measurements, their shifts and, consequently, the variations in $\bar{E}$ extracted from equation 2.

![Figure 8](image.png)

*Figure 8. The influence of surface contamination on the resonance frequency of a 93 nm thick silicon cantilever.*
The shift in frequencies is caused by the combination of both the additional adsorbed mass and the stiffening effect on the cantilever [39]. Therefore, the true stiffening effects of the contamination cannot be obtained by simple subtraction of the shifts in the resonance frequencies. The results from Table 1 clearly show that the experimental measurements of the scale-dependent $\dot{E}$ via the resonance frequency method lead to a considerable number of uncertainties.

<table>
<thead>
<tr>
<th>L×w×t (µm$^3$)</th>
<th>$f_{bv}$</th>
<th>$f_v$</th>
<th>$f_{sv}$</th>
<th>$\Delta f = f_{sv} - f_{bv}$</th>
<th>$\Delta \dot{E}$ (GPa)</th>
<th>$\Delta \dot{E}/\dot{E}$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>96 × 8 × 1.019</td>
<td>123.65</td>
<td>147.968</td>
<td>128.85</td>
<td>5.2</td>
<td>0.80</td>
<td>8.4</td>
</tr>
<tr>
<td>24 × 8 × 0.340</td>
<td>583.73</td>
<td>607.108</td>
<td>585.48</td>
<td>1.75</td>
<td>0.30</td>
<td>0.6</td>
</tr>
<tr>
<td>8 × 8 × 0.093</td>
<td>994.53</td>
<td>1090</td>
<td>973.6</td>
<td>-20.93</td>
<td>-3.90</td>
<td>-4.2</td>
</tr>
</tbody>
</table>

Table 1. The resonance frequency of cantilevers in air and vacuum and the variations in extracted $\dot{E}$. The resonance frequency data are in KHz.

Another popular method which is used to characterize the scale-dependence is the bending tests with the use of force spectroscopy in a scanning probe microscope (SPM) [11]. SPM has very high force and displacement resolution and sensitivity, yet it has significant uncertainties in its measurement and interpretations. Measuring the absolute deflection requires a painstaking calibration of the deflection sensor, as well as the risk of tip slippage [24] and indentation of the probe tip [41]. Another important factor which degrades the accuracy of the method is the lack of knowledge in the precise force. In a typical SPM measurements, the force can only be obtained from a calibrated cantilever with known spring constant. The spring constant of an SPM cantilever usually deviates from its designed nominal value by a significant amount. The lack of an accurate spring constant calibration method, together with the above issues result in inaccuracies which can amount to more than 26% error [11].

In order to avoid the aforementioned issues in experimental investigation of the scale-dependence, we implemented the recently developed measurement method, which uses the quasi-static electrostatic pull-in instability (EPI) phenomenon [32]. The distinctive advantage of the EPI lies in its well-known sharp instability and the possibility of applying an SI-traceable force along the length of the beam [32]. The details of the phenomenon can be found elsewhere [44]. Due to the quasi-static nature of the method, it does not suffer from mass-loading effects. The beauty of the methods lies also in its simplicity and ease of use. The required actuation voltage can be precisely measured using standard electrical test equipment and a microscope. Fig. 9 shows a schematic illustration of the EPI setup. It consists of a voltage source and a semiconductor testing probe station. The pull-in voltage was measured by slowly increasing the voltage difference between the cantilever and the substrate. Visual observation of the color changes and the sudden snap-in of the cantilever were used to determine the pull-in voltage.
For each cantilever, the pull-in voltage and geometry of various cantilevers with different lengths were measured. Fig. 10 shows the experimental measured pull-in voltages versus the length of cantilevers with different thicknesses. By solving the nonlinear differential equation of the electromechanical system, knowing the pull-in voltage and the geometry of the cantilever, one can extract $\hat{E}$ for each thickness [32]. Fig. 11 shows the extracted $\hat{E}$ as a function of the cantilever thickness. Error bars were calculated according to [32]. The maximum calculated error due to uncertainties in the cantilever’s geometry measurements and the measured voltage was 12% [32]. The main error contributions comes from the geometry, and specifically, the length. Results in Fig. 11 show a considerable scale-dependent behavior in the effective stiffness for bending, $\hat{E}$. 

Figure 9. Schematic view of the measurement setup for EPI. The driving voltage on the cantilever is applied through a probe contact and the substrate is grounded. Reprinted with permission from [32]. Copyright 2010 American Institute of Physics.
Above 150 nm the effective modulus is constant and converges to the bulk value of [110] silicon elastic modulus. Somewhere near 150 nm it starts to decrease monotonically. It can be seen that at 40 nm, the $E$ is about half of the bulk value. So the question to be answered is: what causes this scale-dependent behavior?
4. Theoretical investigations: Impact of surface effects and native oxide layers

The influence of surface effects at scales higher than micrometers on the overall elastic behavior is negligible and the elastic behavior can be explained by the bulk properties. However, in nano systems, due to their small size, the surface-to-volume ratio is quite large and therefore,
the surface effects cannot be neglected. This has motivated major research to try to explain the scale-dependence by the surface effects [41, 47, 20, 17]. The surface effects can be categorized into two types; the extrinsic effects such as surface contamination, native oxide layers and environmental adsorptions, and the intrinsic effects which originate from the difference in the atomic configurations near or at the surface and in the core of the bulk. The latter are known as surface stress and surface elasticity effects. The energy associated with the atoms near the surface is different from the energy of atoms in the core. The surface atoms have lower coordination numbers and electron densities and therefore tend to adopt equilibrium lattice spacing differently from the bulk ones [48, 23]. On the other hand, the epitaxial relationship from bulk to surface has to be maintained, therefore, bulk atoms strain the atoms near the surface and create the so-called surface stress [13, 23, 36, 49]. Consequently, the surface atoms like to reconstruct and can deviate easily from their original ideal situations, and therefore, have different elasticity compared to the bulk, known as surface elasticity [41, 47, 20, 23].

In order to predict the surface effects on $\hat{E}$, we have developed a semi-continuum, two-dimensional plane-stress framework [23], which takes into account the influence of surface elasticity on the elastic behavior of nanocantilevers. One of the main modifications to the previous models [47, 20] is the inclusion of coupling stiffness that arises from a surface stress and surface elasticity imbalance caused by either surface reconstruction [50] or molecular adsorption [51] at the top and bottom surfaces [22]. According to this framework, the bending mode of $\hat{E}$ can be calculated as

$$\hat{E} = E_b \left( 1 + \frac{3\Sigma S}{E_b t} \right)$$

(3)

where $E_b$ is the bulk value of Young’s modulus, known to be 169 GPa for [110] mono-crystalline silicon, $\Sigma S$ is the sum of the surface elasticity at the top and the bottom surfaces and $t$ is the thickness of the cantilever. In order to calculate $\hat{E}$, the contribution of the surface elasticity ($\Sigma S$) is required. Because the surface region is only a few atomic layers thick, atomistic calculations are necessarily involved in order to calculate the contribution of the surface elasticity, and in general to include the surface effects in the modeling of nanosystems. We carried out molecular dynamics calculations (MDC) using a recently developed modified embedded atom method (MEAM) potential [52]. Silicon nanobeams and plates were created with initial atomic positions corresponding to the bulk diamond-cubic crystal. The simulation cells were then fully relaxed to a minimum energy state at room temperature and zero pressure. After relaxation, surfaces, such as the one shown in Fig. 12, display a (2×1)–type reconstruction. The details of MDC can be found in [49]. After reaching the equilibrium configurations, the nanoplates were subjected to a quasi-static loading and were again fully relaxed after each strain increment. $\hat{E}$ in extensional mode was defined as

$$\hat{E} = \left( \frac{1}{V_0} \right) \frac{d^2U}{de^2}$$

(4)
where $V_0$ is the volume of the nanoplate in fully relaxed zero-strain configuration (undeformed configuration), $U$ is the total energy of the system and $\varepsilon$ is the applied strain.

The value of $\hat{E}$ was calculated from total energy for simulations with various thicknesses. The results are shown in Fig. 14 (triangles). In order to determine the $\sum S$ using MD, similar to the calculation of $\hat{E}$ described above, the nanoplate is subject to a quasi-static loading and is fully relaxed after each strain increment. The $\sum S$ is extracted in extensional mode as the 2nd-order derivative of surface energy with respect to the strain, and is about -1 Nm$^{-1}$ [29, 49]. This value was used as an input for Equation 3. Here, it is assumed that the surface elasticity in bending and extensional modes are very close [16]. Considering only this surface elasticity effect, a scale-dependence can be calculated and the results are plotted in Fig. 14 as a solid line. The thickness is shown on a logarithmic scale in order to visualize clearly the difference between the experimental results and theoretical investigations. The results of the semi-continuum approach indicate that the surface effects are significant and influence the scale-dependence below 15 nm. It is therefore clear that surface effects can only partially explain the observed scale-dependence.

The native oxide, as an extrinsic surface effect, influences $\hat{E}$ via both its own elastic response and the complex interactions between the oxide and the silicon at the interface [29]. However, it has been shown that the influence of substantial elastic behavior differences between silicon and native oxide is much more significant than the interface elasticity of Si-SiO$_2$. Moreover, during the oxidation process, for every thickness unit of silicon oxide 0.44 units of the silicon surface is consumed, resulting in reducing the thickness of silicon and, consequently, decreasing $\hat{E}$. By taking into account the surface elasticity of the original silicon and the elastic modulus of native oxide layers $E_{Ox}$, $\hat{E}$ can be estimated as [29]
where $E_{Ox}$ and $t_{Ox}$ are the native oxide Young’s modulus and layers’ thickness at the top and bottom of the cantilever. The thickness of the native oxide layers have been reported to vary between 2 to 5 nm [11] and its Young’s modulus reported between 50 to 75 GPa [11, 29]. These variations, especially the thickness (since it has a higher influence on the effective elasticity) have been used to calculate the $\hat{E}$ of silicon cantilevers with different thicknesses. The result is shown in Fig. 13. The dotted and dashed lines in Fig. 14 show a comparison between the resultant $\hat{E}$ as a function of cantilever thickness considering the various native oxide scenarios’ influence (ranges of thicknesses and Young’s modulus reported in the literature) and the measured values. It is clear from the figure that taking the native oxide layer scenarios into account reduces the difference and partially explains the distinctive reduction in $\hat{E}$, yet there is still a considerable gap between experiments and theory.

\[
\hat{E} = \frac{E_s t^3 + E_{Ox} \left(1 + \frac{3 \Sigma S}{E_s t}\right) \left[8\left(\frac{t_{Ox}}{t}\right)^3 + 6t^2 t_{Ox} + 12t \left(\frac{t_{Ox}}{t}\right)^2\right]}{\left(t + 2t_{Ox}\right)^3}
\]  

(5)

Figure 13. The effective Young’s modulus of silicon cantilever as a function of native oxide layers thickness, which varies between 2 to 5 nm and the thickness of the silicon cantilever.
One of the most important issues, which has not previously been taken into account, is considering the fact that experimentally tested nanosystems are not defect-free in contrast to the perfect single crystal structures studied using atomistic simulations. In the next section, we discuss the existence of defects and their influence on the scale-dependence.

5. Impact of fabrication-induced defects

There are still some remaining issues that require resolution in order to close the gap between experiments and theory [25]. Among them, accounting for the fact that experimentally fabricated nanosystems are not defect-free, has been recognized as the most important effect, which has not yet been investigated and quantified [25].

As described in Section 2, reducing the cantilever thickness was done by alternating thermal oxidation and etching in HF solution. It was observed that the density of fabrication-induced defects increased when reducing the thickness of the silicon device layer of SOI wafers. During the thermal oxidation process, the lattices of silicon and silicon oxide mismatch, thus introducing stress in both layers [53]. The precipitated oxygen and local strains in Si-SiO\textsubscript{2} interface are likely sources of defects [29]. Further reduction of the thickness of the silicon layer would result in the initiation of defects and cracks, which can extend through the whole layer.
thickness similar to a pinhole, reaching the underlying BOX layer [53, 29]. These are called HF defects, since the existence of these defects is tested and observed by using the HF etching [38]. The HF solution that reaches the underlying BOX layer etches isotropically the SiO$_2$ underneath, resulting in a suspended membrane-like layer with a tiny defect in the middle. An optical microscope image of a HF defect at a 14 nm thick silicon device layer is shown in Fig. 15.a. The figure clearly shows a buckled membrane with a HF defect in the middle. The buckling is due to the residual stress between silicon and silicon oxide.

The existence of HF defects was also confirmed by SEM, which is observed as a circle (membrane) with the defect at the center (Fig. 15.b). Actually, the HF modifies the defect size and shape, and makes it visible enough to locate it. Using AFM the defect can be observed without HF modifications. The AFM image is shown in Fig. 16.a.

![Figure 15.](image1.png)  
(a) Optical microscope image of a HF defect. (b) SEM of a HF defects. The BOX layer below the HF defects is etched by the HF solution, resulting in the formation of a membrane with a defect in the middle.

![Figure 16.](image2.png)  
(a) AFM image of a defect without HF modification (shape is caused by a AFM probe). (b) Extending the HF etching of the defects for a longer time causes suspension of relatively large membranes, which finally collapse to the layer below and create a Swiss cheese-like structures.
Continuing the HF etching of the defects results in formation of membranes with a much larger diameter, which are no longer stable and collapse to the layer underneath. This has been shown in Fig. 16.b. What is ultimately remaining, is a structure with big holes, similar to a “Swiss cheese” shape.

The appearance of defects on the surface of the cantilevers can be observed with white light interferometry. As an example, Fig. 17.a shows a white light interferometry image of 40 nm thick silicon cantilevers with defects. The depth profile of one of the defects on the cantilever is shown in Fig. 17.b.

![Image of 3D measurements and profile measurement](image)

**Figure 17.** (a) Reconstructed 3D measurements of 40 nm thick silicon cantilevers with defects, using a white light interferometer. (b) The profile measurement of a defect taken along the line 1 in (a), showing the depth and geometry of the defect.

It can be speculated that HF defects contribute to a decrease in $\tilde{E}$ and thus, play a role in the gap between experiments and theory of scale-dependence. We utilized MD to examine the extent that the defects contribute to the scale-dependent behavior. Similarly to Section 4, silicon nanoplates were created along the [100] and [110] directions. Defects with cylindrical shapes of different sizes were created by removing the atoms inside the cylinder. Fig. 18 shows the relaxed [100] and [110] nanoplates with defects. After reaching equilibrium configurations, the nanoplates underwent a quasi-static tensile loading. The virial stresses [54] were calculated for each level of strain and $\tilde{E}$ was calculated from the slope of the resulting stress-strain curve. Details of the calculations were published in [38]. The MD calculations were repeated for different nanoplates (small, medium and large size) with different dimensions of the defects. The ratio of effective elastic modulus with and without defects, $E_d / E_0$, versus the ratio of the defects’ surface area and total surface area, $A_d / A_0$, are plotted in Fig. 19. The inset shows the results as a function of the ratio of the number of removed atoms and total number of atoms, $N_d / N_T$. It is noteworthy that creating the defect did not change the average total energy per atom significantly, except for the atoms near the surface of the defects, which have higher excess energy. An example is shown in Fig. 18.e. The coloring is according to the energy. Apart from the two layers closest to the defect, the energy of the other atoms did not substantially change.
Figure 18. Snapshots of (a) a [100] silicon nanoplate with one defect at the centre, (b) and (c) patterned defects in a [100] silicon nanoplate and (d) [110] nanoplate with one defect at the centre. (e) a mid-plane of a [100] nanoplate with a defect in the centre, indicating the atoms with different energy due to the presence of a defect. Coloring is according to energy.

The reason that the first two layers have a different energy is because removing atoms creates free surfaces inside the defects, causing atoms inside the defect to reconstruct resulting in changes in energy. In order to determine the contribution of each atom to the reduction of $\hat{E}$, the changes in the virial stress of each individual atom as a function of strain, which is the stiffness contribution of each atom, was calculated. The results are shown in Fig. 20, indicating substantial changes when the size of the defects is increased.
Figure 19. Reduction of the effective Young’s modulus versus the size of a defect for different dimensions and directions of nanoplates. The inset shows the re-plotting with respect to the number of removed atoms.

Besides MD, an analytical solution based on continuum theory was used to approximate the effects of defects on $\hat{E}$. It is based on the approximation described in [55] and is as follows:

$$\hat{E}_d = \frac{E_0 \left(1 - \frac{A_d}{A_0}\right)}{1 + 2 \frac{A_d}{A_0}}$$

The result is plotted in Fig. 19 as a dashed line. A remarkably good agreement between MD and analytical solutions can be seen. The results show that $\hat{E}$ decreases monotonically with increasing defect density (defect volume fraction). The decrease is significant and therefore, it has to be considered as an additional factor contributing to the scale-dependence.
Figure 20. Snapshots of middle layers of silicon (a) [110] nanoplate without defect, (b) [110] nanoplate with a defect in the centre, (c) and (d) [100] nanoplates with patterned defects. Each individual atom is coloured by its stiffness contribution. Blue colours denote atoms of higher stiffness.

The experimental observations, which were described earlier in this section, indicate that the defect density \( (A_d/A_0) \) is between 0.005% to 0.12%, depending on the thickness of the silicon device layer [38]. Taking into account the range of experimentally observed defect density and the results presented in Fig. 19, one can include the effect of defects in the scale-dependent \( \hat{E} \). The results including the effects of surface elasticity, native oxide layers and fabrication-induced defects were calculated for the thicknesses that were experimentally measured. The results are presented in Fig. 21. Due to variations in the defects’ density, an upper and lower limit of the influence of defects is introduced in the figure. The shaded area in the figure demonstrates the extent that defects influence \( \hat{E} \). Further quantification of the defects’ density versus the thickness of nanocantilevers has to be done. It can be seen, that taking the defects into account, could explain the observed gap between experimental measurements and theoretical calculations.
Figure 21. Scale-dependent $E$ in bending mode, taking into account the surface elasticity, native oxide layers and the fabrication-induced defects. In order to take into account the variations in the defect density, an upper and a lower limit are introduced. Thus, the scale-dependence observed in silicon nanostructures can be explained by a contribution of surface effects, native oxide layers and fabrication-induced defects.

6. Summary, conclusions and recommendations

The main observations, investigations and discussions to be drawn from the chapter are summarized and concluded in this section.

Silicon based nanodevices are widely used in sensing and actuating applications. For reliable design of such devices, a thorough knowledge of the mechanical properties of these nanostructures is of vital importance. In order to improve the sensitivity, significant research efforts have been directed towards reducing the size of the nanostructures. However, decreasing the size causes the mechanical behavior and the elastic behavior to deviate from the bulk value, known as “scale-dependence” phenomenon.
Two approaches have been explored to study scale-dependent elastic behavior: experimental and theoretical, however a discrepancy exists between the two approaches. The scale at which the scale-dependence starts in experimental measurements is different from that estimated theoretically. Various techniques involving resonance frequency, tensile tests in TEM, atomic force microscopy and nanoindentation have been used to characterize the effective Young’s modulus of silicon nanocantilevers and nanowires. Each method involves different assumptions, sources of errors and interpretations. We proposed the use of the electrostatic pull-in instability method in order to avoid most of the issues with other existing methods, as described in Section 3. Using molecular dynamics calculations and the semi-continuum approaches (Section 4) the surface effects on the elastic behavior of silicon nanocantilevers have been investigated. Direct comparison between the surface effects simulations with experimental data from the reliable EPI method showed that although surface effects influence the effective Young’s modulus of silicon to some extent, they alone are insufficient to explain the experimental observations.

Another important influence is caused by native oxide layers that exist at the surfaces of the silicon nanocantilevers. The native oxide layers influence the effective Young’s modulus of silicon in 3 ways. Through: 1) its distinct elastic response; 2) unknown interactions between the oxide and the silicon at the interface; and, 3) consumption of silicon during oxidation; during oxidation for every unit of silicon oxide 0.44 units of the silicon surface is “consumed”. We assumed that the effect of native oxide on the surface elasticity of silicon, or the interface elasticity of Si-SiO2, is not significant compared to the distinct elastic response of the native oxide. Taking the native oxide layers into consideration reduced the difference between experimental measurements and theoretical predictions, yet there was still a considerable difference (Fig. 14).

One of the important issues which was confirmed experimentally, but had not previously been considered in theoretical modeling, was accounting for the fact that experimentally tested nanocantilevers and nanowires are not defect free. Molecular dynamics calculations have been carried out to determine the effects of defects on the effective Young’s modulus of silicon nanocantilevers. The conclusion is that the scale effect observed in silicon nanocantilevers can be explained by a contribution of surface effects, native oxide layers and defects. Taking these into account, the gap between the experimental measurements and theory can be closed.

There are a number of issues that are recommended for future research on the size effects:

- The electrostatic pull-in instability method has its own limitations; the application of electrostatic load requires a fairly conductive device and a counter electrode. This implies that the method is only applicable to conductive materials. Moreover, stiction prevention by additional stopper is necessary to ensure the release of the cantilever for multiple measurements. Another fact, that is believed to be minor, but has to be investigated in detail, is the influence of the electric field (electrostatic charges) used in EPI method on the measured value of the effective Young’s modulus of nanocantilevers and nanowires.

- Future work on the silicon native oxide layer and its properties will help to explain in more detail the observed size effects. The scale-dependent elastic behavior of silicon native oxide
layers has to be investigated. Proper measurements on determination of the thickness and the Young’s modulus of silicon native oxide layer as a function of silicon thickness is recommended as further research. In this work, we assumed that the effect of native oxide on the surface elasticity of silicon, or the interface elasticity of Si-SiO$_2$ is not significant. However, a systematic computational study on the interface of Si-SiO$_2$ is essential to determine the extent to which this factor affects the elastic behavior. One of the limiting factors for studying the interface of Si-SiO$_2$ is the lack of a proper potential for Si-SiO$_2$ in MD calculations. Therefore, future research for developing a reliable potential for silicon native oxide and its interface with silicon is also necessary to enable direct comparison with experiments.

In order to better explain the influence of HF defects on the effective Young’s modulus of silicon nanosystems, future work on the quantitative determination of the defects and defect density as a function of silicon thickness is necessary. Moreover, the inevitable random distribution of defects, arising from nanocantilever and nanowire synthesis has to be taken into account.

**Acknowledgements**

This book chapter was financially supported by Enabling Technology program (ETP) for Materials Technology, Dr. L.J.M.G. Dortmans, of Netherlands Organization for Scientific Applied Research (TNO). The authors would like to thank Dr. T.S.D. O’Mahoney and Dr. A. Bossche for their supports, helps and valuable suggestions.

**Author details**

Hamed Sadeghian$^{1,2}$, Fred van Keulen$^2$ and Hans Goosen$^2$

1 Technical Sciences, Netherlands Organization for Applied Scientific Research, TNO, CK, Delft, The Netherlands

2 Structural Optimizations and Mechanics Group, Department of Precision and Microsystems Engineering, Delft University of Technology, CD, Delft, The Netherlands

**References**


