Development of Bistable Out-of-Plane SMA Nanoactuators

Zur Erlangung des akademischen Grades eines

Doktors der Ingenieurwissenschaften (Dr.-Ing.)

von der KIT-Fakultät für Maschinenbau des
Karlsruher Instituts für Technologie (KIT)

genehmigte

Dissertation

von

M.Eng. Zixiong Li

Tag der mündlichen Prüfung: 28 Juli 2025
Hauptreferent: Prof. Dr. Manfred Kohl
Korreferent: Prof. Dr.-Ing. Alfred Ludwig

Abstract

Having the advantages of large work density and potential scability of dimensions down to nanometer size, shape memory alloys (SMAs) are competitive candidates among the active materials for applications in micro- and nanoelectromechanical systems (MEMS and NEMS). Being compatible with standard semiconductor fabrication techniques, SMAs can be designed as an active layer on a common substrate such as silicon (Si). With at least these two layers, such systems can convert input energy into work if constructed as freely movable and mechanically active structures, being the base elements of micro- and nanoactuators.

This work focuses on the development of SMA-based nanoactuators to generate work in out-of-plane direction with the overall objective to perform bistable actuation at the nano scale enabling novel smart functionalities in optics and photonics. Potential applications are, for instance, mechanically tunable photonic directional couplers using SMA nanoactuators to control light transmission in silicon waveguide structures. Joule-heated multilayer beam structures are the common shapes of SMA-based nanoactuators. A mechanical and a thermal analytical model are developed that can conveniently predict the mechanical deformation and temperature profile, respectively, as well as electrical power consumption of such multilayer beam structures when being Joule-heated. The thermal model reveals that heat conduction to the substrate dominates the heat dissipation mechanism at micrometer-scale. The accuracy of the thermal model is validated experimentally.

Starting from the analyses of the fundamental SMA/Si bilayer cantilever beams, this work extends the investigations of SMA-based nanoactuators to the developments of bistable SMA nanoactuators and SMA-based mechanically tuned photonic directional couplers. Appropriate fabrication processes are developed for the designed devices accordingly, which include the techniques of electron beam lithography (EBL), reactive ion etching (RIE), physical vapor deposition (PVD), etc. $Ti_{54}Ni_{31}Cu_{15}$ is selected as the SMA material used for the developments of nanoactuators in this work. In particular, the magnetron sputtering of $Ti_{54}Ni_{31}Cu_{15}$ films and the stress engineering of the films are conducted within the frame of close collaborations with the project partners from Ruhr University Bochum (RUB).

The temperature-dependent electrical resistance measurement confirms that the 600 nm $Ti_{54}Ni_{31}Cu_{15}$ film on Si substrate transforms above room temperature with a narrow hysteresis width of 1.9 °C. Revealed by the collaborating partners from RUB, the transformation property is independent of film thickness down to 150 nm. On this basis, the electro-thermo-mechanical performances of the Joule-heated $Ti_{54}Ni_{31}Cu_{15}$/Si bilayer cantilever nanoactuators are studied analytically and numerically. The mechanical analyses

show that the deflection change can be attributed to competitions among different effects: mismatch of thermal expansion coefficients, variation of elastic moduli (ΔE) and shape memory effect. There is an optimal thickness ratio of 0.5 between SMA and Si causing the maximum stroke. The thermal analyses define a range of heating current that 80% of beam length completes phase transformation while overheating is avoided. The partially transformed part of the beam is folded into a wing-structure to reduce the temperature gradient. The joint electro-thermo-mechanical analyses prove that the optimal range of the wing ratio (the length ratio of the folded wing to the beam) is 20%–25% to trade off among the stroke, power consumption and structural stiffness.

Bistable structures are famous for the appealing properties of low energy consumption and non-linear force-displacement relation. A design of bistable out-of-plane SMA nanoactuators is proposed, which consists of four parallel arranged cantilever beams. By exploiting the residual stress in the $Ti_{54}Ni_{31}Cu_{15}$ film, the center beams are deposited with $Ti_{54}Ni_{31}Cu_{15}$ film to introduce buckling of the two outside beams. The two stable buckling states are switched by utilizing the effect of mismatch of thermal expansion coefficients and the shape memory effect. In the fabrication process containing two lithography steps, the $Ti_{54}Ni_{31}Cu_{15}$ films are selectively deposited on free-standing structures by lift-off with success. The initial deformations of the fabricated nanoactuators indicate that there is considerable position-dependent internal stress in the Si layer of the center beams, which hinders the bistability by increasing the moment threshold of state-switching. Nanoindention test of the nanoactuator shows a snapping process during loading with a stroke about 3.6% of beam length, which is accompanied by a surge of force about 98.8 μN. This non-linear force-displacement relation demonstrates that extra energy input is necessary to snap from one stable buckling state to the other. In addition, the linear part of the force-displacement curve prior to the non-linear part discloses that the moment from residual stress should be overcome in advance to enter the snapping stage.

In the field of integrated photonics, MEMS-tuning exhibits distinct advantages in terms of small footprint and high energy efficiency. The previously developed $Ti_{54}Ni_{31}Cu_{15}$/Si bilayer cantilever nanoactuator is integrated with photonic directional coupler to mechanically tune the coupling between the waveguide structures. Because of the movable structures, particular attention is paid to the design of photonic circuits, which utilize taper structures and both rib waveguides and strip waveguides to minimize the loss when switching the bottom claddings between SiO_2 and air. Working at TE-mode, the directional coupler is designed for the full coupling at the length and gap size of 36.6 μm and 110 nm, respectively, which is proven to have negligible influence on the mechanical deflection of the nanoactuator. In the fabrication process containing three lithography steps, the utilization of rib waveguides provides another advantage of being a robust hard mask for the wet etching step of SiO_2, which simplifies the fabrication by avoiding additional materials as etching mask. Transmission tests of reference waveguides confirm that the tapers in the photonic circuits and the annealing process of SMA have negligible effect on the transmission loss. Due to the initial deformation of the free-standing directional couplers, 71% of light power is coupled at the designed full coupling length. Because another test with fixed directional couplers validates the numeric model of directional

coupler, it is highlighted that precise stress controlling and anisotropy of SMA sputtering are critical for the optimal tuning performance.

Zusammenfassung

Aufgrund ihrer großen Arbeitsdichte und der möglichen Skalierbarkeit der Abmessungen bis in den Nanometerbereich sind Formgedächtnislegierungen (FGL) vielversprechende Kandidaten unter den aktiven Materialien für Anwendungen in mikro- und nanoelektromechanischen Systemen (MEMS und NEMS). Da sie mit den Standard-Halbleiter-Herstellungstechniken kompatibel sind, können FGL als aktive Schicht auf einem gewöhnlichen Substrat wie Silizium (Si) hergestellt werden. Mit mindestens diesen beiden Schichten können solche Systeme die zugeführte Energie in Arbeit umwandeln, wenn sie als frei bewegliche und mechanisch aktive Strukturen konstruiert werden, die die Grundelemente von Mikro- und Nanoaktuatoren bilden.

Diese Arbeit konzentriert sich auf die Entwicklung von FGL-basierten Nanoaktuatoren zur Erzeugung von Arbeit außerhalb der Substratebene mit dem übergeordneten Ziel, eine bistabile Aktuierung auf der Nanoskala zu ermöglichen und damit neue intelligente Funktionen in Optik und Photonik zu realisieren. Mögliche Anwendungen sind zum Beispiel mechanisch abstimmbare photonische Richtkoppler, die FGL-Nanoaktuatoren nutzen, um die Lichtübertragung in Silizium-Wellenleiterstrukturen zu steuern. Durch Joule'sches Beheizen aktuierte Mehrschicht-Balkenstrukturen sind eine gängige Bauform von FGL-basierten Nanoaktuatoren. Es werden ein mechanisches und ein thermisches analytisches Modell entwickelt, mit dem sich die mechanische Verformung, das Temperaturprofil sowie der elektrische Energieverbrauch der Mehrschicht-Balkenstrukturen bei Joule'schem Beheizen vorhersagen lassen. Das thermische Modell zeigt, dass die Wärmeleitung zum Substrat den dominierenden Mechanismus der Wärmeabfuhr im Mikrometerbereich darstellt. Die Genauigkeit des thermischen Modells wird experimentell bestätigt.

Ausgehend von der Analyse grundlegender FGL/Si-Bimorph-Biegebalken werden in dieser Arbeit die Untersuchungen von FGL-basierten Nanoaktuatoren auf die Entwicklung von bistabilen FGL-Nanoaktuatoren und mechanisch abgestimmten photonischen Richtkopplern auf FGL-Basis erweitert. Für die entworfenen Bauelemente werden geeignete Herstellungsverfahren entwickelt, darunter Techniken wie Elektronenstrahllithographie (EBL), reaktives Ionenätzen (RIE) und physikalische Gasphasenabscheidung (PVD). Als FGL-Material für die Entwicklung der Nanoaktuatoren wird in dieser Arbeit $Ti_{54}Ni_{31}Cu_{15}$ ausgewählt. Insbesondere das Magnetron-Sputtern von $Ti_{54}Ni_{31}Cu_{15}$-Schichten und die Spannungsoptimierung der Schichten erfolgen in enger Zusammenarbeit mit Projektpartnern der Ruhr-Universität Bochum (RUB).

Messungen des temperaturabhängigen elektrischen Widerstands bestätigen, dass die 600 nm dicke $Ti_{54}Ni_{31}Cu_{15}$-Schicht auf einem Si-Substrat oberhalb Raumtemperatur mit einer schmalen Hysteresebreite von 1,9 °C umwandelt. Wie von den Kooperationspartnern der

RUB aufgezeigt wurde, ist diese Umwandlungseigenschaft unabhängig von der Schichtdicke bis zu 150 nm. Auf dieser Grundlage werden die elektro-thermo-mechanischen Eigenschaften der mit Joule'schem Beheizen aktuierten $Ti_{54}Ni_{31}Cu_{15}$/Si-Bimorph-Biegebalken-Nanoaktuatoren analytisch und numerisch untersucht. Die mechanischen Analysen zeigen, dass die Änderung der Durchbiegung auf den Wettbewerb verschiedener Effekte zurückgeführt werden kann: die differentielle Wärmeausdehnung, die unterschiedlichen elastischen Module (ΔE-Effekt) sowie der Formgedächtniseffekt. Es gibt ein optimales Dickenverhältnis von 0,5 zwischen der FGL-Schicht und dem Si-Träger, bei dem ein maximaler Hub auftritt. Die thermischen Analysen legen einen Bereich des Heizstroms fest, bei dem 80% der Balkenlänge eine vollständige Phasenumwandlung durchlaufen und zugleich eine Überhitzung vermieden wird. Der nur teilweise umgewandelte Teil des Biegebalkens wird zu einer Flügelstruktur gefaltet, um den Temperaturgradienten zu verringern. Die gemeinsamen elektro-thermo-mechanischen Analysen zeigen, dass ein Flügelverhältnis (Längenverhältnis des gefalteten Flügels zum Balken) von 20%–25% den besten Kompromiss zwischen Auslenkung, Energieverbrauch und struktureller Steifigkeit darstellt.

Bistabile Strukturen sind bekannt für ihre attraktiven Eigenschaften wie niedrigen Energieverbrauch und nichtlineare Kraft-Auslenkungs-Kennlinien. Es wird ein Entwurf für bistabile SMA-Nanoaktuatoren zur Bewegungserzeugung aus der Ebene vorgeschlagen, der aus vier parallel angeordneten Balkenträgern besteht. Unter Ausnutzung der Eigenspannung in der $Ti_{54}Ni_{31}Cu_{15}$-Schicht werden die mittleren Balken mit $Ti_{54}Ni_{31}Cu_{15}$ beschichtet, um ein Ausknicken der beiden äußeren Balken zu erreichen. Die beiden stabilen Knickzustände werden durch Ausnutzung der unterschiedlichen Wärmeausdehnungskoeffizienten und des Formgedächtniseffekts umgeschaltet. Im Herstellungsprozess mit zwei Lithografieschritten werden die $Ti_{54}Ni_{31}Cu_{15}$-Schichten selektiv auf freistehenden Strukturen durch Lift-Off mit Erfolg abgeschieden. Die anfänglichen Verformungen der hergestellten Nanoaktuatoren deuten darauf hin, dass in der Si-Schicht der mittleren Balken erhebliche positionsabhängige Eigenspannungen vorhanden sind, die die Bistabilität erschweren, indem sie die Schwelle zum Umschalten der Zustände erhöhen. Ein Nanoindentationstest des Nanoaktuators zeigt einen Schnappvorgang während der Belastung mit einem Hub von etwa 3,6% der Balkenlänge, der von einem Kraftstoß von etwa 98,8 µN begleitet wird. Diese nichtlineare Kraft-Weg-Beziehung zeigt, dass zusätzliche Energie notwendig ist, um von einem stabilen Knickzustand in den anderen zu schnappen. Außerdem zeigt der lineare Teil der Kraft-Weg-Kurve vor dem nichtlinearen Teil, dass zunächst das Moment der Eigenspannung überwunden werden muss, um in den Schnappbereich zu gelangen.

Im Bereich der integrierten Photonik bietet das Abstimmen mit MEMS-Technologie deutliche Vorteile hinsichtlich geringer Baugröße und hoher Energieeffizienz. Der zuvor entwickelte $Ti_{54}Ni_{31}Cu_{15}$/Si-Bimorph-Biegebalken-Nanoaktuator wird mit einem photonischen Richtkoppler integriert, um die Kopplung zwischen den Wellenleiterstrukturen mechanisch abzustimmen. Aufgrund der beweglichen Strukturen wird besonderes Augenmerk auf das Design der photonischen Schaltkreise gelegt, die Taperstrukturen sowie Rippen- und Streifenwellenleiter nutzen, um die Verluste beim Wechsel der unteren Schicht zwischen SiO_2 und Luft zu minimieren. Der für den TE-Modus ausgelegte Richtkoppler ist für eine vollständige Kopplung bei einer Länge von 36,6 µm und einem Spalt von 110 nm ausgelegt, was nachweislich einen vernachlässigbaren Einfluss auf die mechanische Ablenkung des

Nanoaktuators hat. Im Herstellungsprozess, der drei Lithografieschritte umfasst, bietet die Verwendung von Rippenwellenleitern den zusätzlichen Vorteil, als robuste Hartmaske für den Nassätzschritt von SiO_2 zu dienen, was die Fertigung vereinfacht, da keine zusätzlichen Materialien als Ätzmaske benötigt werden. Transmissionstests an Referenzwellenleitern bestätigen, dass die Taperstrukturen in den photonischen Schaltkreisen und der Temperprozess der FGL einen vernachlässigbaren Einfluss auf die Transmissionsverluste haben. Aufgrund der anfänglichen Verformung der freistehenden Richtkoppler werden 71% der Lichtleistung bei der vorgesehenen vollen Kopplungslänge eingekoppelt. Da ein weiterer Test mit feststehenden Richtkopplern das numerische Modell des Richtkopplers validiert, wird betont, dass eine präzise Spannungskontrolle und Anisotropie der FGL-Sputterschicht entscheidend für eine optimale Abstimmleistung sind.

Contents

List of Figures

List of Tables

List of Main Symbols

(EI)	Bending stiffness
(GJ)	Torsional stiffness
α	Coefficient of thermal expansion
β	Propagation constant
δ	Deflection of beam
κ	Coupling coefficient of directional coupler
λ	Wavelength (in Chapter 3 and Chapter 7) Thermal conductivity (in Chapter 5)
σ_c	Electrical conductivity
σ_{re}	Mechanical residual stress
ρ	Electrical resistivity
θ	Excess temperature
φ	Curvature of curve
ϕ	Bending angle of beam
ξ	Volume fraction of martensite phase
A_s	Austenitic start temperature
A_f	Austenitic finish temperature
E	Elastic modulus
F_{cr}	Critical force of buckling
$^{1\mathrm{st},\ 2\mathrm{nd}}$	(Superscript) First and second mode of buckling
F_{eq}	Equivalent force

L_c Coupling length of directional coupler

M_{eq} Equivalent bending moment

M_s Martensitic start temperature

M_f Martensitic finish temperature

n Refractive index

$\boldsymbol{B}$ Magnetic flux density vector

$\boldsymbol{D}$ Electric displacement vector

$\boldsymbol{E}$ Electrical field

$\boldsymbol{H}$ Magnetic field

$\boldsymbol{J}$ Current density vector

$\boldsymbol{k}$ Wave vector

$\boldsymbol{S}$ Poynting vector

1. Introduction

1.1. Background

1.1.1. Smart Materials

For centuries, scientists and engineers have been dedicated to improving the properties of existing materials and creating new materials to meet the demands of the fast developing technologies. In the past decades, a new concept "smart materials" has emerged. By taking advantages of multifunctionalities, this category of materials can respond to environmental stimuli (temperature, light intensity, humidity, variation of pH, electric and magnetic fields, etc.) by changing their shapes or physical properties [1–3]. Table 1.1 lists some of the multifunctionalities among different physical inputs and outputs. Materials showing these effects provide the fundamental basis for the actuators and sensors of the next generation.

Table 1.1.: Some multifunctionalities of the smart materials among different inputs and outputs.

Input \ Output	Mechanical	Electrical	Magnetic	Thermal	Optic
Mechanical		Piezoelectric, Piezoresistive	Inverse magnetostrictive	Elastocaloric	Photoelastic
Electrical	Electrostrictive, Inverse piezoelectric, ER fluid[1]		Maxwell-Ampère Law	Electrocaloric, Peltier effect	Electrochromic, Electroluminescence
Magnetic	Magnetostrictive	Magnetoresistive		Magnetocaloric	Magneto-optic
Thermal	Thermo-mechanical	Thermoelectric[2], pyroelectric	Thermomagnetic		Thermochromic
Optic	Optical gradient force	Photoelectric	Inverse Faraday effect	Photothermal	Photochromic, Fluorescent

Shape memory materials, which encompass alloys and ceramics, are a well-studied category of smart materials. Their appealing properties are based on the phase transformations under external stimuli with variations of their microstructure. Most of the shape memory

[1] Electrorheological fluid.
[2] Including the Seebeck effect and the Thomson effect.

alloys (SMAs) are thermally responsive that generate deformations at changing temperatures (thermomechanical effect). Examples of such alloy systems include Nitinol (Ni-Ti alloy) and Ni-Ti based ternary alloys, copper (Cu)-based Cu – Zn – Al and Cu – Al – Ni alloys [4]. A special subgroup are ferromagnetic SMAs that exhibit deformations induced by the magnetic field (magnetostritive effect) [5]. An example of this type is the Ni – Mn – Ga alloy [6]. In addition, because of the entropy change during phase transformations, the SMAs show a temperature variation when being deformed in adiabatic environment (elastocaloric effect) [7]. Owing to their distinct material properties, the versatile SMAs have been widely exploited in different engineering fields, which covers aerospace [2, 8], biomedical [9–11] and robotics [12], etc. In particular, the implementations in microelectromechanical systems (MEMS) are of special interest. It has been extensively explored to integrate SMAs into microactuators due to the large work density up to $10^7\,\mathrm{J/m^3}$, which is 1–2 orders higher than other common actuators [4, 13]. Successful examples include microgrippers [14, 15], microvalves [15] and nanoactuators [16]. Other properties of SMAs are also well employed, for instance, in the applications of microcooling devices [17] and energy harvesters [18].

The phenomenon of thermally induced phase transformation is also witnessed in certain ceramics, such as zirconia (ZrO_2) and vanadium dioxide (VO_2) [19, 20]. Notably, besides the mechanical deformation, the phase transformation of VO_2 is also accompanied by considerable changes of both electrical resistance and refractive index (thermochromic effect) [21]. Hence, this transition of physical properties is also known as metal-insulator transition, which takes place at the temperature of 68 °C. Apart from the mechanical utilization in micro- and nanoactuators [20–23], VO_2 films are also widely used for light tuning and sensing purposes. At macroscale, VO_2 films are explored in the design of smart windows that are transparent to infrared light at room temperature while become opaque at elevated temperature [24, 25]. At microscale, the material is adopted for the design of on-chip photonic modulators, where the phase transition of VO_2 can be triggered by various physical stimuli, comprising thermal [26–29], electrical [30, 31] and optic [32].

1.1.2. Bistable Structures

Bistable structures are mechanical systems or devices that possess two distinct and stable equilibrium states. These configurations are characterized by their ability to remain stable in either state without the continuous application of external forces or energy input. Transition between the two states (snap-through) usually occurs through a suitable external trigger, which overcomes the energy barrier between the stable states and causes the structure to snap from one stable state to the other. The underlying principle of bistability is typically rooted in the nonlinearity, which originates from the material properties and the structural design [33].

Material nonlinearity is attributed to the non-linear stress-strain relation of the material. For example, SMAs are well-known for the hysteresis of strain change during either mechanical or thermal loading cycles. This property was exploited in designing a bistable structure made by two different types of SMAs with different hysteresis widths [34, 35].

The bistability that presents in the SMA-polymer composite structures is another example of utilizing the material nonlinearity [34, 36, 37]. Polymers experience a sharp change of elastic moduli by three orders of magnitude during the glass transition [38, 39], which is used to capture either the martensitic or austenitic shape of the SMA layer by controlling the highest temperatures in different thermal cycles.

Structural nonlinearity refers to the non-linear load-displacement relation stemming from specially designed structures instead of materials, and the commonly adopted method is to introduce buckling of the structure. Based on the classical Eulerian buckling theory, the differential deformation equation naturally has multi-solutions corresponding to different stable states. Over the past decades, different approaches have been investigated to create buckling. Applying an external compressive force on a slender beam in the axial direction is a direct and simple way to achieve buckling [40, 41]. However, to maintain the buckling shape, the compressive force is required to be continuously applied. Designing the structural geometry in a pre-deflected shape has been proved to be a more practical approach and has received substantial attention from the scientific community [42–47]. Recently, there has been increasing interest in tuning the internal stress in thin films to obtain the buckled structures. In a structure with three beams arranged in parallel, Pan et al. selectively introduced the residual stress to the outside two beams with pulsed laser so that the middle beam could buckle [48]. Similar structures are also witnessed in [49, 50] while fabricated at μm scale, and the stress for buckling comes from the residual stress of deposition and thermal stress. Meanwhile, some unconventional structures showing bistability have also been proposed. For instance, Mao et al. designed a latching bistable microswitch consisting of two bilayer electrothermal beams. The latching state can be switched by controlling the beam that is heated [51].

The stimuli to switch between the stable states are not limited to mechanical loading. By taking advantage of different physical principles, diverse switching mechanisms have been explored. Electric, magnetic and temperature fields are commonly used for generating non-contact stimuli [33]. Besides, some unconventional strategies have been explored as well. For instance, Song et al. utilized the pressure from phase change of fluid to trigger the state switching [43]. Dong et al. designed an optomechanical circuit that the buckling state of a doubly-clamped bridge could be switched by the optical gradient force from a neighboring ring resonator [52]. These different switching approaches effectively extend the applicability of the bistable structures.

Having the unique advantages in energy efficiency, mechanical simplicity and down-scaling capability, bistable structures have drawn growing attention not only in the macro-scale domain, but also in the field of micro- and nanoelectromechanical systems (MEMS and NEMS). As a kind of movable structures, integrating the bistable structures into the design of micro- and nanoactuators is a well-investigated application field [33]. Moreover, research on other potential applications in various domains have been reported, including switching [42, 51], memory storage [52, 53] and energy harvesting [40, 54].

Nevertheless, bistable structures also present challenges. Due to the nature of nonlinearity, the designs are usually non-intuitive, and there are strong interactions among different parameters. During the design phase, precisely controlling the structural performances

usually requires sophisticated material constitutive equations and structural deformation theories and models. As an illustration, different analytical models have been proposed to depict the buckling and state switching process of the classical pre-deflected structures, which vary with the loading and boundary conditions [44, 47, 55–59].

1.1.3. MEMS for Integrated Photonics

The term "integrated photonics" refers to the integration of photonic components on a common substrate [60]. These components include waveguides, power splitters, polarizers, modulators, etc. Similar to the electronic integrated circuits (IC), these photonic components can be carefully arranged to build photonic integrated circuits (PIC) for achieving certain functions, e.g., for the applications in high-speed telecommunication [61]. Benefiting from the large difference of refractive indices between silicon (Si) and silicon dioxide (SiO_2) [62], the "silicon-on-insulator" (SOI) wafers, which are commonly used in the IC industry [63], are adopted for the fabrication of PIC. The compatibility of the fabrication techniques and processes of PIC with the highly sophisticated IC industry underlines another attractive point of the development of PIC.

Efficient tuning mechanisms, which change the optical properties of waveguides, are essential for the successful applications of PIC. Selecting an appropriate tuning method depends on the material and the tuning purpose. For the Si-based SOI platform, effective tuning methods include thermo-optic (temperature-dependent refractive index of Si), Kerr effect (an electro-optic effect), plasma dispersion (Si belongs to semiconductor materials), optical gradient force (optomechanics) and MEMS [64, 65]. Among these methods, MEMS tuning possesses the advantages of low power consumption, low loss, relatively small footprint and independence from waveguide materials [64, 66]. For instance, integrating with bistable structures introduced in the previous section can significantly reduce the power consumption of tuning. Nevertheless, the implementation of MEMS tuning is hindered by limitations of several aspects. Usually, free-standing structures bring additional challenges in fabrication and are fragile in nature. Moreover, the tuning frequency is limited by the inertia of movable devices to the mechanical resonance frequency [64], which is generally lower than the electro-based tuning mechanism.

Because light is an oscillating electromagnetic wave propagating in the medium, the basics of tuning are the modifications of either the amplitude or the phase of the wave. In practical, these two factors can be mechanically influenced through the following approaches [64]:

- Varying the refractive index (either real or imaginary part) of the waveguide material by stress / strain.
- Varying the propagation length by displacement or stress.
- Placing a specific material block within the field of the waveguide mode (displacement variation).

Based on the three principles, Fig. 1.1 demonstrates the detailed mechanical tuning mechanisms for several purposes, which include phase shifting, power splitting, switching and tuning of the grating couplers.

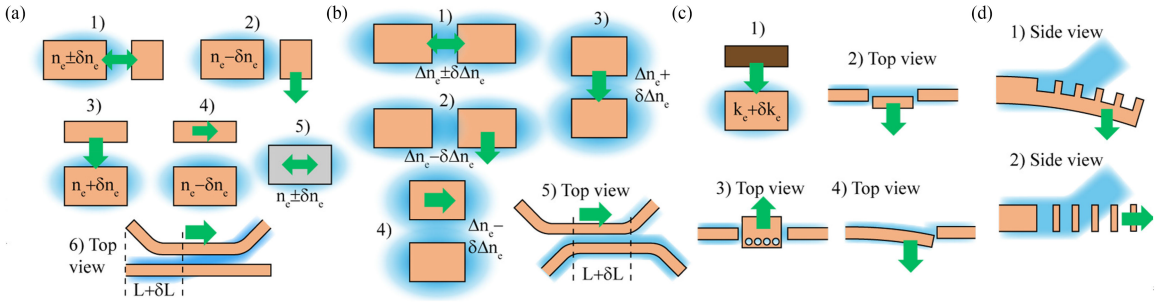


Figure 1.1.: Schematics of several MEMS tuning mechanisms. (a) Phase shifting by varying the effective refractive index with loading 1) in-plane, 2) out-of-plane, 3) double-layer out-of-plane, 4) double-layer in-plane, 5) strain tuning, 6) waveguide elongation. (b) Power splitting by variations 1) in-plane, 2) out-of-plane, 3) double-layer out-of-plane, 4) double-layer in-plane, 5) in longitudinal direction. (c) Switching by 1) placing an absorbing material, 2) displacing a waveguide section, 3) inserting a photonic crystal, 4) bending the waveguide. (d) Tunable grating couplers by 1) bending a grating, 2) varying the grating period with in-plane actuation. Picture adapted from [64].

So far, different photonic components tuned by MEMS have been reported. A widely studied category is the directional coupler used for power splitting[3]. The coupler is actuated either in-plane by a comb-drive actuator (mechanism (b1) in Fig. 1.1) [67, 68] or out-of-plane by a bilayer cantilever (mechanism (b2)) [69]. Moreover, one of the waveguides in the directional coupler can be specially designed in the shape of a buckled bistable beam, which lies on top of the other (mechanism (b3)) [70]. The state switching is achieved by the comb-drive actuator providing the switching force in vertical direction. Another component that has been mechanically tuned is the ring resonator. By applying a voltage between the ring and the substrate, the ring deforms under the electrostatic force, functioning as a reconfigurable wavelength filter (mechanism (a6)) [71]. Reversely, by exploiting the strain-dependent refractive index (mechanism (a1)), the ring resonator was designed as a strain gauge which is able to measure the external loading by the shifting of resonating wavelength [72].

Due to the nature of free-standing structures, the main drawbacks of MEMS-tuned PIC come from the on-chip cavities where the movable structures are embedded. The perturbations from the environment, such as humidity and contamination particles, can easily influence the performance of the movable structures. Therefore, reliable hermetic packaging techniques are essential [73, 74]. Another impediment is the mode mismatch at the edge of the cavities that causes transmission loss. This mismatch comes from the sudden change of the cladding materials. Appropriately designed taper structures are feasible solutions to minimizing the loss [75]. In addition, tapers are also beneficial for improving the mechanical strength at the fixed-end of free-standing structures.

[3] When the splitting is 100%, the device also functions as switch.

1.2. Outline

By leveraging the large work output of SMAs, this dissertation aims at the development of bistable out-of-plane SMA nanoactuators and the potential utilization in integrated photonics for tuning purpose. The outline of this dissertation is as follows:

Chapter 2 gives a brief introduction to two aspects. The first section reviews the phenomenon, the variation of microstructures and the characterization method of the phase transformation property of SMAs. The second section summarizes the techniques involved in the fabrication processes of SMA-based nanoactuators.

Chapter 3 outlines the essential background of the integrated photonics, including fundamentals of the guided mode in the waveguide and the basic functioning principles of the photonic components that are utilized in this dissertation.

Chapter 4 derives the analytical models that can efficiently give the mechanical and thermal performances of the multilayer beam structures, which are the fundamental structures of SMA nanoactuators. The wing structure design has been integrated into SMA nanoactuators to homogenize the temperature gradient, and special attention is paid to the mechanical stiffness relaxation brought by this design.

Chapter 5 analyzes the mechanical, thermal and joint thermo-mechanical performances of SMA/Si bilayer nanoactuators. To mitigate the temperature gradient induced by Joule-heating, the wing structure is integrated into the nanoactuator, and the corresponding optimal range of wing ratio is investigated to compromise among the mechanical stroke, electric power consumption and structural stiffness.

Chapter 6 develops a design of bistable out-of-plane nanoactuators by exploiting the residual stress in the SMA layer. This introduces buckling to the beam structure and leads to two stable states. Geometrical parameters, including the lengths and widths of the beams, are studied numerically to ensure the appropriate critical buckling force leading to the first buckling mode. As demonstrators, the bistable SMA nanoactuators are fabricated and characterized by nanoindentaion tests.

Chapter 7 integrates the bilayer SMA nanoactuators with the photonic directional couplers, which are free-standing and can be actively tuned. Particular attention is paid to the design of photonic components and the development of a fabrication process that is compatible to free-standing directional couplers and nanoactuators. Although the functionalities of the fabricated nanoactuators are hindered by technical problems, transmission tests are conducted to characterize the optical performances of the directional couplers.

2. Shape Memory Alloys and Nanofabrication Techniques

This chapter introduces the basics and fundamentals of shape memory alloys and nanofabrication techniques, which provide the foundations for following chapters. In the first section, the introduction covers the crystallography of shape memory alloys and the martensitic transformation that underlines the well-known shape memory effect. Afterwards, the focus shifts to thin film shape memory alloys: the fabrication, size-effect and the characterization methods. The next section outlines the techniques that are involved in the fabrication of nanoactuators, including lithography, pattern transferring and etching.

2.1. Shape Memory Alloys

2.1.1. The Material

Shape memory alloys (SMAs) are a unique group of smart materials that can generate work output when the microstructure is changed under specific external stimuli, such as heat and magnetic field [5]. Being responsive to external heat, thermal SMAs have two different phases depending on the temperature: the highly symmetric austenite (or parent) phase typically with body-centered cubic crystal structure (B2) at high temperature, and the martensite phase with either tetragonal, orthorhombic or monoclinic crystal structure at low temperature [2, 5]. The exact crystal structure depends on different factors, such as composition. Ni-Ti (Nickel-Titanium) alloys possess a monoclinic structure (B19') while adding element such as copper (Cu) or palladium (Pd) changes the lattice to orthorhombic (B19) or creates an intermediate rhombohedral R-phase [4]. As an example, Fig. 2.1 illustrates the sketches of the crystal structure and microstructure of Ni-Ti alloy at martensite, austenite and R-phase, respectively.

Because of the non-symmetric structure of martensite, this phase has different orientations (variants) with respect to the parent phase. Kumar and Lagoudas have proven that there are 24 martensitic variants in the Ni-Ti alloy system in total [4]. It can be classified into two categories by the stress level, i.e., the twinned martensite under low stress and detwinned martensite under high stress. The former is composed of self-accommodated variants, and the latter is constituted by variants with a specific type in dominant, which is favored by the orientation of external stress.

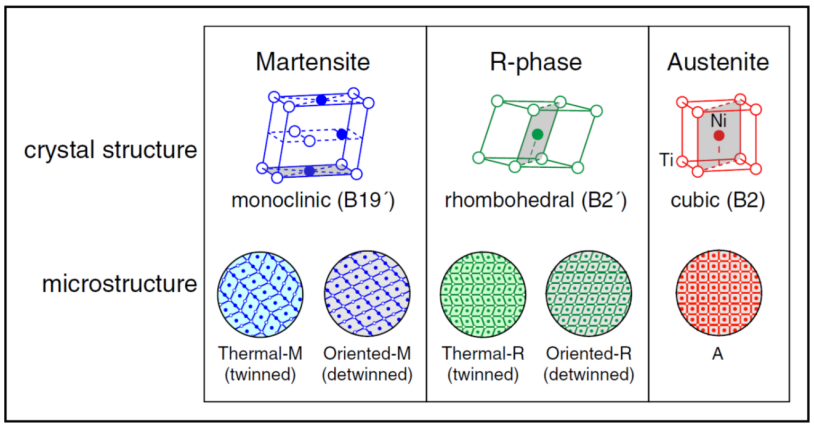


Figure 2.1.: Schematics of the crystal structure and microstructure of Ni-Ti alloy at different phases. Diagram adapted from [76].

The martensitic transformation, i.e., from the parent phase to martensite, is witnessed in various alloy systems such as Fe-C and Cu-Zn [77, 78]. However, to have the transformation in a reverse direction, i.e., from martensite to the parent phase, it is crucial that there is no plastic deformation in the martensitic transformation [79]. The microscopic explanation is that there is no newly created lattice defect to have reversible transformation [77]. The mechanical stress induced during martensitic transformation can be compensated either by dislocation slip, which is irreversible, or by twinning and self-accommodation processes, which is reversible [77]. The martensitic transformation in SMAs belongs to the latter situation, therefore, it is also named as *thermoelastic* martensitic transformation.

The martensitic and its reverse transformations (also termed as *forward* and *reverse* transformation) underline the special effects of SMAs and the potential applications. Belonging to the first-order solid state phase transformation [5, 76], the martensitic transformation is diffusionless due to its nature of twinning, i.e., a shear distortion of lattice structure, hence the compositions of the austenitic and martensitic phases are the same. During the martensitic transformation, within a single crystal the martensite forms along the habit planes, which remain neither distorted nor rotated, thus these planes are known as invariant planes. Consequently, the invariant planes are also the interfaces between austenite and martensite.

The martensitic transformation shows a temperature or stress hysteresis at a given stress or temperature level. Correspondingly, at stress-free state, the transformation can be characterized by four important temperatures, i.e., M_s, M_f and A_s, A_f. The first two refers to the temperatures at which the forward transformation starts and finishes, respectively, and the last two refers to the temperatures at which the reverse transformation starts and finishes, respectively. An example of such temperature hysteresis curve is presented in Section 5.2.

Under certain conditions, the intermediate R-phase exists in Ni-Ti alloys during the forward transformation so that the phase variation sequence is austenite-R-martensite. This phase emerges by elongating along any one of the <111> direction of the B2 crystal structure of

parent phase [79]. It has been concluded that there are basically three ways to introduce the R-phase: (1) cold-work, (2) aging of the alloys with higher Ni content and (3) the addition of another element [4, 5]. The mechanism is to introduce stress fields into the crystal structures by dislocations (from cold-work or aging), precipitations such as Ni_4Ti_3 and Ni_3Ti_2 in Ni-rich Ni-Ti alloys, and third element [4, 77]. Despite the advantage of narrow and stable hysteresis in austenite-R transformation, the low transformation strain (around 1%) [5] limits the range of actuation applications.

Among different SMA systems, the Ni-Ti system has been extensively studied and put into industrial usage [5, 9]. In particular, some research focuses on introducing a third element that can tune the transformation temperatures and the hysteresis width to meet specific requirements. For example, iron (Fe) and chromium (Cr) decrease transformation temperatures [5], whereas hafnium (Hf) can significantly increase transformation temperatures to above 100 °C [37, 80, 81]. For the applications of actuation, the Ni-Ti-Cu ternary system, in which Cu partially replaces Ni in the composition, is advantageous compared with alloys containing other substituting elements. Adding Cu reduces the hysteresis width [4, 5, 9], which is beneficial to fast response and having small working temperature range. Besides, the Ni-Ti-Cu system shows less sensitivity of the transformation temperatures to composition compared to the Ni-Ti binary system [4, 82], and the transformation temperatures can be above room temperature for actuation applications [83].

2.1.2. Shape Memory Effect

SMAs have two pronounced effects, namely the shape memory effect (SME) and pseudoelastic effect (PE). The thermomechanical paths of these two effects are illustrated in Fig. 2.2, in which SME follows the path 1-6 and PE follows the path 7-13. The major difference between these two effects is the temperature that the SMA material is deformed:

- If the deformation takes place at the low-temperature martensitic state, the material can recover to its original shape by heating above A_f. This phenomenon is SME.
- If the deformation takes place at the high-temperature austenitic state, the material can recover from large strain up to 8% [2, 5], to its original shape by removing the load isothermally. This phenomenon is PE.

The property of SME that the SMA material can generate large strain by heating forms the basis of applying SMAs for actuation [5, 77], and PE is usually used in passive applications such as damping [5]. Therefore, the introduction in the following only focuses on the SME.

Point 1 to Point 2: At Point 1 of the low-stress and low temperature condition, the material is at twinned martensite state. Loading to Point 2 follows the typical elastic deformation of metals before yielding.

Point 2 to Point 3: The stress reaches the critical value that the twinned martensite begins to reorient and forms the preferred detwinned martensite. The macroscopic

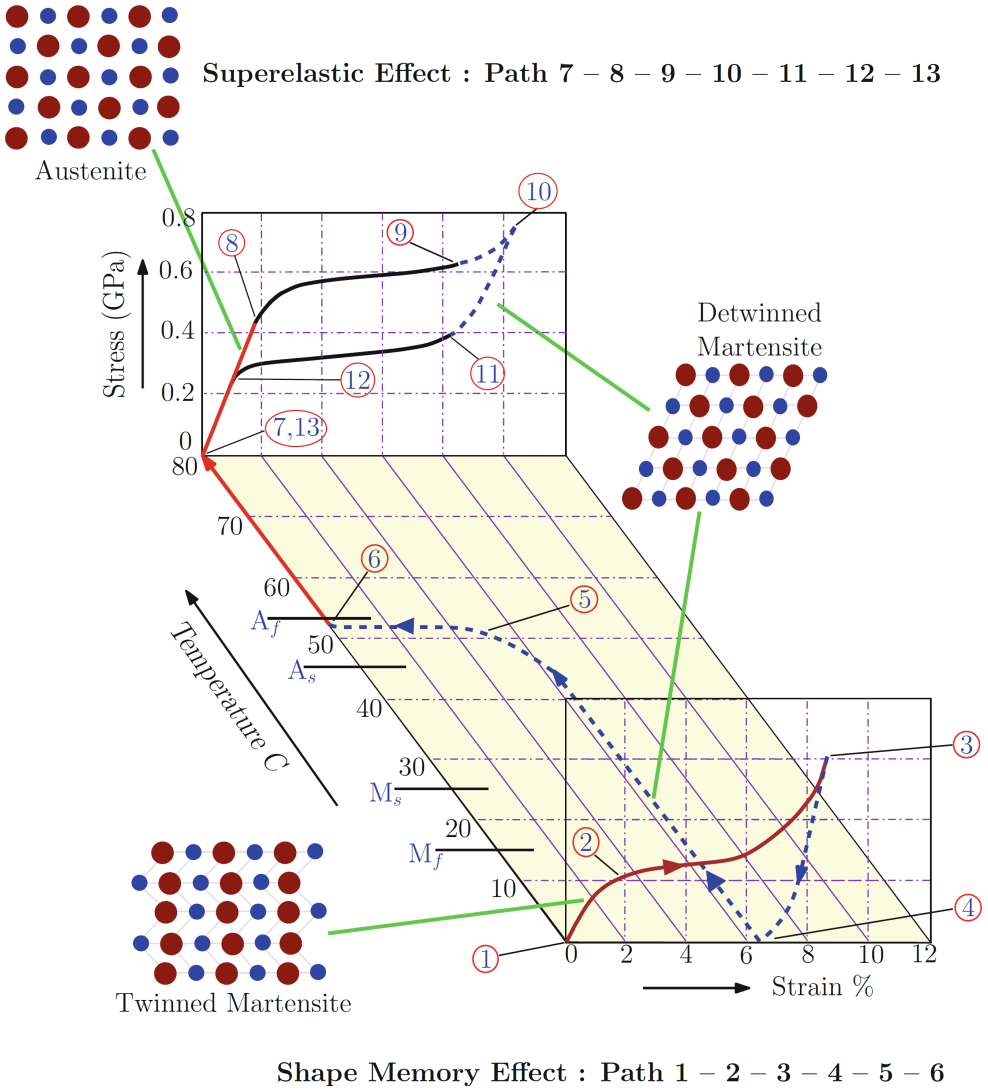


Figure 2.2.: A three-dimensional strain-stress-temperature graph illustrating both the shape memory effect (path 1–6) and the pseudoelastic effect (path 7–13) of SMAs. Picture adapted from [76].

deformation corresponds to a plateau stage of the stress-strain curve, which ends when the detwinning transition completes. Further loading to Point 3 leads to elastic deformation again of the detwinned martensite.

Point 3 to Point 4: Unloading from Point 3 follows a path of elastic unloading until reaching Point 4, where the material is still at detwinned martensite with a retained strain of about 4–6% [2].

Point 4 to Point 6: The material is heated up and the reverse transformation completes when the temperature is above A_f. The residual strain is fully recovered during the transformation from the twinned martensite to austenite. The macroscopic phenomenon is that the deformed SMA material at low temperature returns to its original "memorized" undeformed shape by heating.

Point 6 to Point 1: In the absence of external stress, the SMA material undergoes the forward transformation back to the twinned martensite. There is no macroscopic strain change during the transformation.

The description above is also named as *one-way SME* as only the shape at high temperature is "memorized". If the SMA material has experienced the training of certain thermomechanical loading cycles, it can also exhibit *two-way SME* that the material can repeatably switch between the shapes of martensitic and austenitic states respectively in thermal cycles without applying an external mechanical load. The training process stabilizes the hysteresis responses of SMAs and saturates the inelastic strain. The mechanism behind training is that training introduces permanent defects in the crystal structures, which create internal residual stress and lead to the formation of the preferred detwinned martensite [4]. Thus, variation of internal stress state may disturb the two-way SME of SMAs.

2.1.3. SMA Thin Films

Conventionally, the bulk SMA-based components at macroscale are fabricated based on casting or powder metallurgy techniques [10], whereas deposition techniques are the widely used methods to prepare SMA thin films with thickness below 5 μm. Usually, the purely physical process, direct current (DC) magnetron sputtering, is preferred, which has higher deposition rate than the radio frequency (RF) sputtering [84]. Other techniques, such as electron beam evaporation, pulsed laser deposition and chemical vapor deposition (CVD), have also been reported in research [9, 84].

The general process of DC magnetron sputtering is described as follows. In a high vacuum chamber, the ions of inert gas (normally Ar) are accelerated to impinge on the metal targets, causing metal atoms to be knocked out from the targets and build up on the substrate. For Ti-Ni-Cu film deposition, this process is conducted on a cold substrate. Therefore, the deposited film is amorphous. A post-annealing process is required to crystallize the film and render the phase transformation property. In addition, it is favorable to firstly deposit a thin interfacial layer, such as Cr or tantalum (Ta), with thickness of about 20 nm before sputtering the SMA layer, which is helpful to promote adhesion and prevent delamination [85].

Unlike bulk SMA, because the thickness of thin film is usually substantially smaller than the in-plane dimensions, the small thickness can have an influence on the material properties of SMAs, especially when it is comparable to the grain size. For instance, a surface oxide layer (TiO_2) can form on Ni-Ti based SMA thin films when exposed to air, which is responsible for the formation of surface wrinkling and trenches of the films [86]. Moreover, this oxide layer not only introduces additional constraints on the lattice during phase transformation, but also creates a Ti-depleted zone underneath and shifts the composition of SMA. The phenomenological influence is that the transformation temperatures decrease whereas the hysteresis width increases [87]. Similarly, the diffusion at the interface between the SMA layer and the Si substrate can also significantly change the composition of SMA, hence disturb the phase transformation [88]. Introducing a diffusion stop layer, for instance

SiO_2, is beneficial to ameliorate this degradation, which has been verified in the study of the substrate-dependent critical thickness of Ti-Ni-Hf films [37]. This thickness effect becomes ignorable when the thickness is larger than the average grain size [87], and the material properties of SMA thin films become similar to those of bulk material.

2.1.4. Characterization of Phase Transformation Temperatures

Determining the transformation temperatures is one of the preliminary and fundamental characterization tests for a material with phase transformation properties. In terms of the thermally induced SMAs, the differential scanning calorimetry (DSC) testing is a widely used thermal analysis, which effectively reveals the transformation temperatures and the corresponding latent heat [89, 90]. The principle behind is that, as a first-order phase transformation, the enthalpy is discontinuous during the transformation, i.e., there is change of the latent heat. In detail, the forward transformation is an exothermic process (heat emitting) whereas the reverse transformation is an endothermic process (heat absorbing) [89]. However, when the investigated SMA is in the form of sputtered thin film on a wafer substrate, and the film thickness is below 1 µm, preparing free-standing film samples for DSC testing is challenging. The temperature-dependent electrical resistance measurement is a more common and convenient method to identify the transformation temperatures of SMA thin films [89, 90]. The resistance of the sample is monitored by the four-probe setup, as depicted in Fig. 2.3. Four probes are placed equidistantly, and a constant current source is located between the outer two probes. The inner two probes are linked to a voltmeter. In addition, it is assumed that the film thickness t is far smaller than the spacing distance s between the probes ($t \ll s$).

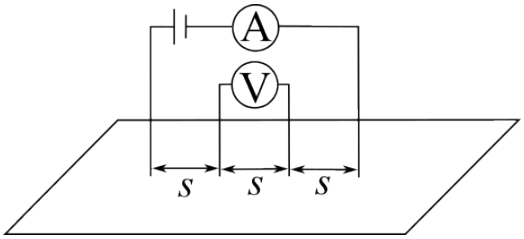


Figure 2.3.: Sketch of the four-point measurement setup for thin film samples.

To record the temperature-dependent resistance data, the setup in Fig. 2.3 is placed on a thermostat inside a vacuum chamber, which can prevent oxidization at high temperature and frosting at low temperature. The ambient temperature in the chamber is controlled under the quasi-stationary condition to guarantee the thermal equilibrium state of the sample at each set temperature point.

Furthermore, the temperature-dependent resistivity of film $\rho(T)$ can also be obtained with the following equation [91]:

$$\rho(T) = \frac{\pi}{\ln 2} \frac{U(T)}{I} t. \tag{2.1}$$

Other characterization techniques, including the temperature-dependent X-ray diffraction (XRD) [89, 90, 92] and the magnetic measurement for characterizing the transformation of magnetic SMAs [89], are beyond the research scope of this dissertation.

2.2. Nanofabrication Techniques

Generally, the common MEMS and NEMS fabrication processes consist of the three basic steps: thin film deposition, lithographic patterning and pattern transfer. Each of these steps corresponds to various methods and techniques depending on the dimensions, material properties and functions of the target devices to fabricate. The following introductions solely focus on techniques that are exploited in this dissertation.

2.2.1. Maskless Lithography

Lithography is usually the first procedure in NEMS fabrication to project the designed patterns onto the resist layers on the wafers. Typically, the patterns are two-dimensional (three-dimensional grayscale lithography [93] is beyond the discussion here) and engraved on masks that can selectively shade the radiation sources such as ultraviolet (UV) light and X-rays. However, mask-required lithography techniques are both time- and cost-consuming for the research and development phase of new devices, during which there can be several iterations of the patterns of the designs. Maskless lithography eliminates the need for pre-prepared masks by using radiation beams and scanning across the wafer surface pixel by pixel. The common maskless lithography techniques include direct laser writing (DLW) and electron beam lithography (EBL). The working principle of the former is similar to the standard photolithography, while the latter is recognized for its high resolution, as indicated by the Rayleigh criterion [94]

$$Resolution = \frac{k_1 \lambda}{NA}, \tag{2.2}$$

where λ is the wavelength and NA is the numerical aperture. The wavelength of the electrons is about 3.7 pm at the acceleration voltage of 100 kV [95], which is significantly smaller than the wavelength of the UV light for exposure, such as i-line of 365 nm. Although the actual resolution of EBL is limited by the electron scattering phenomenon instead of Eq. 2.2, the EBL system still possesses a finer resolution (at the order of 10 nm [96, 97]) compared to UV lithography, particularly in the case of using thin resist layer. Hence, it is eligible to define structures with feature sizes at nanometer scale.

The basic working principle of EBL is using the accelerated electrons to interact with the resists. For positive resists, the long chains of the resists are broken into short chains by the electrons; while for negative resists, the short chains of the resists are cross-linked into long chains. The most common positive resist is poly-methyl-methacrylate (PMMA), which comprises a series of different products distinguished by the relative molecular

mass. PMMA with smaller molecular mass requires lower dose value to expose (higher sensitivity), hence it shows lower resolution and contrast. On the contrary, PMMA with higher molecular mass needs larger exposure doses and has higher resolution and contrast [97]. As for the negative resists, there are also polymer-based resists such as the AR-N 7520 series. A particular type is hydrogen silsesquioxane (HSQ), which cross-links into non-soluble silicon oxide after being exposed [98]. Owing to its higher etching selectivity in comparison with polymers, HSQ is particularly advantageous when the exposed resist is used as the hard mask for the subsequent dry etching step. This advantage is demonstrated in the fabrication process in Chapter 7.

The scattering phenomenon of the electrons is a significant issue to consider, which can lead to infidelity of the exposed patterns and is the determining factor of the resolution of EBL. In general, there are three types of electron scattering [96], which are

Forward scattering: The incident electrons are spread in a certain width range that is larger than designed. Higher acceleration voltage and thinner resist layer weaken the forward scattering.

Backscattering: Electrons reflect at the interface of the resist layer and the substrate, which can expose the resist as well. Higher acceleration voltage and higher density of the substrate lead to more severe backscattering.

Secondary electron emission: The incident electrons with high energy excite the low-energy secondary electrons in the resist and the substrate, which determines the actual resolution limitation of EBL at the order of 10 nm [96, 97]. The secondary electrons also account for the major interactions with the resists.

The electron scattering is the main reason of the so-called "proximity effect". The adjacent areas of the designed patterns are also exposed, which causes overexposure in the areas between the shapes and unbalanced dose value among the edges and the center within a shape. Fig. 2.4 demonstrates an example of the proximity effect. In this SEM picture, the two waveguide patterns are defined by the negative resist. Because of the proximity effect, the vicinities of the waveguides are overexposed thus the gap between the waveguides is filled with the residuals of the resist, which should be clear and clean. A practical way to tackle this problem is to assign different dose values depending on the pattern densities and dimensions over the wafer. The areas with large or compact patterns should be exposed with lower value and for the areas with small and sparse patterns, higher dose values are more desirable. In addition, it is always recommended to have dose tests to find the optimal values before the fabrication exposures.

2.2.1.1. Alignment Marker for EBL

In a fabrication process that requires at least two lithography steps, precise alignments of the patterns of different exposure steps are essential. They rely on the robustness of the high-quality alignment markers predefined on the wafers. Usually, these markers are made of metals with high atomic number Z [99], for instance gold (Au) and platinum (Pt). Based

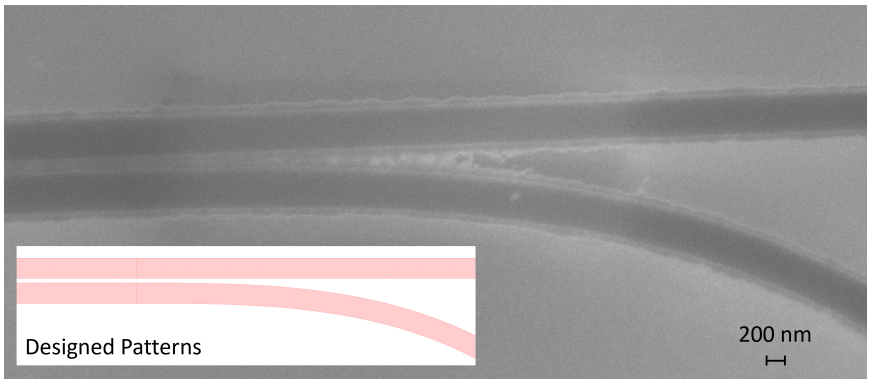
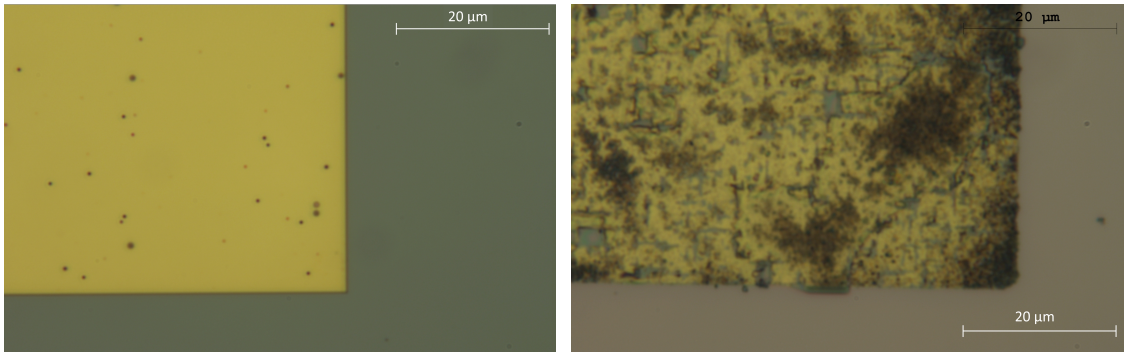


Figure 2.4.: An example showing the proximity effect. In the SEM image, the gap between the two waveguides is filled with the residuals of the negative resist. The bottom-left corner shows the designed patterns from the layout file.

on the Mott scattering [100], higher-Z materials yield enhanced backscattering hence exhibit brighter contrast in the EBL system [97]. However, it is noticed that the metal markers degrade severely and even lose the functionality, if high-temperature annealing or hot-sputtering of the functional films is involved in the fabrication process and this high-temperature step is not the final step. As presented in Fig. 2.5, the evaporated gold layer of the marker is dense before annealing, while it becomes porous after experiencing the annealing step for the SMA layer. It is verified experimentally that the markers with the layer in Fig. 2.5(b) cannot be located correctly in the EBL system. Besides, experiments show that neither having thicker nor thinner layer can ameliorate this degradation, and the same phenomenon is also witnessed with markers made of Pt layer.



(a) Before annealing step. (b) After annealing step.

Figure 2.5.: Comparison of a corner of the gold marker (a) before and (b) after the annealing step for the SMA layer.

A potential interpretation of this observation is the thermal dewetting of the Au and Pt layers on Si substrate [101]. This effect initiates the growth of voids at the grain boundary grooves at the temperature 128 °C for Au and 339 °C for Pt [101], which are considerably lower than the typical annealing temperatures of SMAs, e.g., 600 °C for $Ti_{54}Ni_{31}Cu_{15}$ films. To avoid this degradation phenomenon, groove markers are introduced, which are fabricated by etching into the substrates for a few micrometers [99], allowing to replace

the metal markers in the high-temperature involved fabrication process. The sketch in Fig. 2.6 compares the mechanisms of contrast generation between the two types of markers: the high-Z metal marker creates a bright field by yielding more backscattered electrons, whereas the groove marker creates a dark field by confining the backscattered electrons [102].

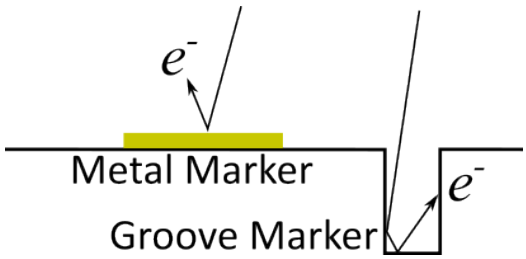


Figure 2.6.: Comparison of the mechanisms between the metal marker and the groove marker.

2.2.2. Additive Pattern Transfer

This section focuses on one of the widely adopted additive pattern transfer techniques, the lift-off process, which is typically performed after the lithography step. Since obtaining high-resolution patterns of metal layers through etching is usually difficult, lift-off process offers an efficient and convenient method for patterning metal films. Fig. 2.7 depicts the typical process flow of the lift-off process. The critical points to obtain successful lift-off patterns are listed below.

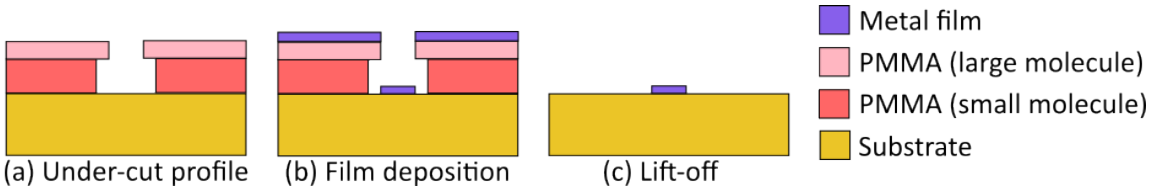


Figure 2.7.: Process flow of lift-off process.

I. The under-cut profile of the resist layer: As shown in Fig. 2.7(a), the exposed resist needs to have an undercut profile to ensure the discontinuity of the deposited films, i.e., the width of the openings of the exposed resists increases in the thickness direction from the top surface. This profile guarantees discontinuity of the deposited metal film without ligaments of top and bottom layers so that the sidewalls of the resist can be reached by the organic solvent in step (c).

However, usually the profile of the exposed resist is overcut, which means that the width on the top surface is wider. A practical approach to have the undercut profile is by making use of the molecular mass-dependent sensitivity of PMMA. As introduced in Section 2.2.1,

when applying the same amount of doses, the PMMA with low molecular mass is more sensitive thus the width of the exposed area after development is larger than the PMMA with high molecular mass. Therefore, a bilayer PMMA with the PMMA of high molecular mass on the top can naturally obtain the undercut profile after exposure [103].

II. The thickness of the resist: The empirical value is that the thickness of the resist layer needs be at least 2–3 times thicker than the thickness of the deposited films. The purpose of having this thickness difference is also to ensure the discontinuity of the metal film. Usually, this requirement defines the upper limit of the film thickness that can be deposited, when developing a fabrication process involving lift-off.

III. Anisotropic film deposition at low temperature: In step (b), it is always preferred to have the deposition in vertical direction, i.e., anisotropically, to ensure discontinuity. Besides, deposition at high temperature should be avoided as the polymer resists experience a glass transition and become soft, causing the deformation of the exposed patterns.

IV. The minimum size of the patterns: In general, the minimum feature size should exceed the grain size of the deposited films to ensure continuous patterns of the films after lift-off process.

It should be highlighted that the lift-off process in Fig. 2.7 usually applies for having the film layer with the area significantly smaller than the area of the wafer, due to the positive resist nature of PMMA. For the situation that the covered area is comparable to the area of the wafer, there is research using a bilayer of PMMA and HSQ to have the "negative tone lift-off" process [104–109], as demonstrated in Fig. 2.8. In step (c), the dry etching of PMMA by using O_2 plasma is isotropic so that the undercut profile is formed naturally. As an example, Fig. 2.9 shows a SEM picture with the HSQ/PMMA bilayer resists.

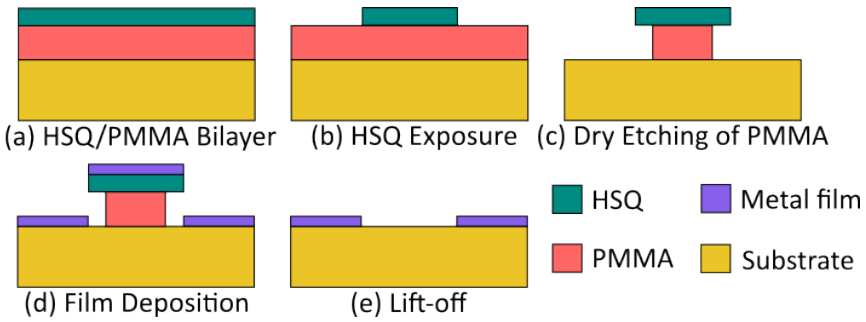


Figure 2.8.: Process flow of the negative tone lift-off process by using the HSQ/PMMA bilayer resists.

2.2.3. Subtractive Pattern Transfer

Contrary to the additive pattern transfer, the subtractive process (etching) refers to selectively removing the materials from the substrate to generate the desired patterns. This

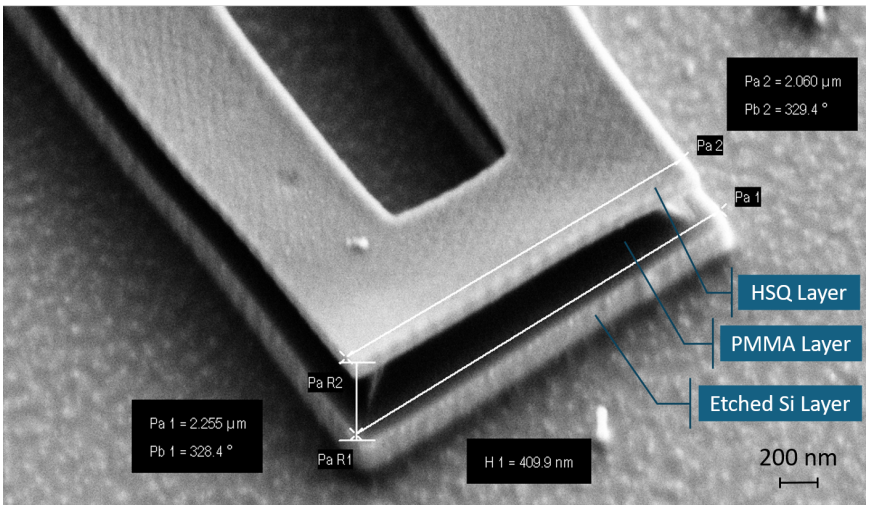


Figure 2.9.: A SEM image showing HSQ/PMMA bilayer resists on a Si substrate. The undercut profile can be clearly seen because of the isotropic etching of PMMA.

process can be a purely physical effect, a purely chemical effect, or a combined effect of both. Two of the basic metrics to characterize the etching process are selectivity and anisotropy. The former indicates the relative etching rates of the etched materials with respect to the mask materials [110]. High selectivity is especially preferred when performing deep etching. The latter indicates the ratio of the etching rates in the vertical direction and the horizontal direction. The etching process can be divided into two categories by the states of the etchant [110]: when it is in the liquid state, the process is named as wet etching; when the etchant is in gaseous or plasma state, it is referred to as dry etching.

2.2.3.1. Reactive Ion Etching (RIE)

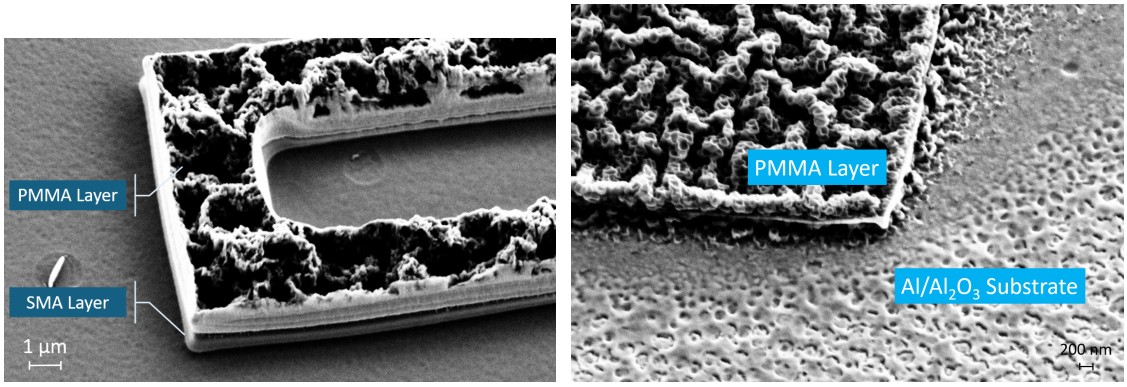
RIE is a plasma-assisted etching process that the reactive gases are ionized into the plasma inside the low-pressure reaction chamber. Ions from the plasma are accelerated and guided towards the target substrate that is connected to an electrode by radio frequency (RF). In general, both physical and chemical reactions take place in the chamber by two types of species [110]: the chemically active neutral radicals and the accelerated ions with bombardment energy. Competition between these two types of species determines the anisotropy of the etched profile on the substrate. There are various parameters that can influence this competition, including the RF power, the flow rates of gases and the substrate temperature [111].

Usually the highly chemically reactive halogen-based gases are used as the etchant, and it is of great importance that the reaction products from the substrate are volatile, i.e., have high vapor pressure, which is beneficial for the continuity of etching. In terms of Si etching, SF_6 is one of the most commonly used gas etchants. The detailed reaction mechanism is complex (54 different reaction formulas have been summarized in [111]). Concisely, the etching is isotropic at room temperature [112], whereas at the temperature below the boiling temperature of SiF_4 at $-86\,^{\circ}C$, the etching becomes anisotropic due to the

formation of the passivation layer (SiO_xF_y) on the sidewalls. In practice, the temperature of this anisotropic etching process is usually set around −100 °C and named as *cryogenic deep RIE* (cryo-DRIE), which is widely used for fabricating Si structures with high aspect ratio [110, 113, 114]. When heating back to room temperature, the passivation layer evaporate away thus there is no residual of the byproduct.

2.2.3.2. Ion Beam Etching (IBE)

IBE is a purely physical plasma-assisted etching process relying on the bombardment of the ions with kinetic energy, which are usually generated by the inert gases such as Ar. Due to the nature of physical processes, IBE shows high anisotropy and low selectivity regardless of the material. Hence, the control of processing time and the selection of the mask material are highly important. Materials with high hardness, for example Cr and Al_2O_3 [115, 116], are more favorable to use as mask. In addition, surface quality is usually poor after IBE. As an example, Fig. 2.10(a) shows a 330 nm-thick SMA layer on the Si substrate completely patterned by IBE. A 4 µm-thick PMMA is used as the mask, which is greatly consumed to about 400 nm after IBE. As a comparison, in Fig. 2.10(b) a 150 nm-thick aluminum (Al) layer is hardly etched in the same process as Al can be easily oxidized to Al_2O_3 in the air naturally.



(a) 330 nm-thick SMA layer after IBE. (b) 150 nm-thick Al layer after IBE.

Figure 2.10.: Comparison of (a) 330 nm-thick SMA layer and (b) 150 nm-thick Al layer after IBE process. 4 µm PMMA masks are heavily consumed in both cases.

2.2.3.3. Wet Etching

Wet etching takes place in a liquid environment, and it is a purely chemical etching process. In terms of the wet etching of SiO_2, it is commonly etched by the hydrofluoric (HF) acid-based solutions with the following reaction [117]

$$SiO_2 + 6\,HF \longrightarrow H_2SiF_6 + 2\,H_2O.$$

The reaction product H_2SiF_6 is water-soluble. Frequently used etchants include the diluted HF (with water) and the buffered oxide etch (BOE), the latter is a mixture of ammonium fluoride (NH_4F) and concentrated HF with a certain ratio. The etching rate is dependent on various factors, including the type of SiO_2, the type and concentration of the etchant solution. Although this etching process is isotropic theoretically, anisotropic etching is witnessed in the experiment, as presented in Fig. 2.11. The BOE wet etching is to release the Si beam structure (the bright area in the figure). Because of the underetching effect, there is remaining SiO_2 under the Si structure. The small slope of the profile of the remaining SiO_2 indicates that the etching rate is higher in the horizontal direction than in the vertical direction.

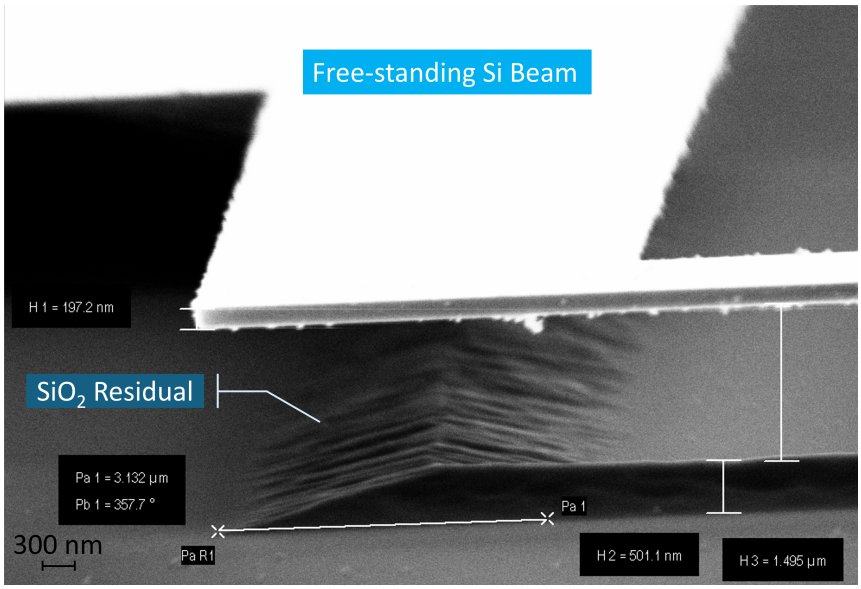


Figure 2.11.: The profile of the remaining SiO_2 showing anisotropy of the wet etching by BOE.

2.2.4. Critical Point Drying

In the fabrication of nanomechanical structures, the SiO_2 layer is commonly used as the sacrificial layer. The wet etching of this layer allows the structural layers on top of it to become movable and also fragile. The stiction caused by the surface tension can potentially damage the free-standing structures when they are dried in conventional method that the liquid is evaporated directly. By avoiding the phase transition from liquid to gas, the critical point drying (CPD) is an effective solution to prevent this damage when removing the liquid in which the structures are rinsed. The fundamental principle is to make use of the properties of the supercritical fluid state that there is no distinction between the gas and the liquid phases, therefore the surface tension also disappears. To reach the supercritical fluid state, both the temperature and the pressure need to be above the critical point [117, 118].

In practice, carbon dioxide (CO_2) is selected as the liquid to perform the CPD, due to its easily reachable critical point of 7.38 MPa at 31.1 °C [117]. Usually, the wafer with released

structures after the wet etching of SiO_2 is rinsed in deionized (DI) water. The etched wafer is transferred to a chamber filled with isopropyl alcohol (IPA), which will be replaced by liquid CO_2 at the room temperature and pressure. IPA is selected as the intermediate liquid because of its excellent miscibility with both DI water and liquid CO_2. Afterwards, the temperature and pressure in the chamber are increased to be above the critical point of CO_2. The temperature is maintained until the CO_2 is completely vented so that the wafer is dried without experiencing the phase transition from liquid to gas.

2.3. Summary

This chapter gives an overview of two aspects. The first aspect is the crystallography of SMAs and the SME of this category of material. Determining four transformation temperatures, i.e., M_s, M_f, A_s, A_f, are important to characterize SMAs. The second aspect is the major techniques in the field of nanofabrication. The designed patterns are firstly exposed on the resist by lithography, then transferred to the substrate by either additive or subtractive approaches. The relative definitions, methods and techniques from this chapter will be cited directly in the following chapters.

3. Integrated Photonics

Integrated photonics refers to the technology that miniaturizes and integrates the photonic components onto a common substrate to perform complex optical functions. By leveraging the fabrication techniques similar to those used in semiconductor electronics, integrated photonics enable compact, high-performance, and cost-effective photonic circuits. Additionally, photonic circuits possess advantages such as low power consumption and low latency in comparison with electronic circuits [119]. Nowadays, there have been continuous advancements of applications in different fields, including telecommunications [61] and quantum technologies [120]. The integration with MEMS has also attracted the interest of researchers aiming for more efficient tuning and modulating mechanisms of integrated photonic devices [64, 66].

Before introducing the work of integration of SMA-based nanoactuator and nanophotonic devices, it is essential to firstly review the basics of integrated photonics and have an introduction of the components that are adopted for the research work.

3.1. Basics of Electromagnetic Waves

3.1.1. Wave Equations

Based on the classic theory, light is a type of electromagnetic radiation propagating in free space or a medium in the form of coupled electric and magnetic fields. This coupling is described by the Maxwell's equations [60, 121] as

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \tag{3.1}$$

$$\nabla \times \boldsymbol{H} = \frac{\partial \boldsymbol{D}}{\partial t} + \boldsymbol{J}, \tag{3.2}$$

$$\nabla \cdot \boldsymbol{B} = 0, \tag{3.3}$$

$$\nabla \cdot \boldsymbol{D} = \rho, \tag{3.4}$$

where $\boldsymbol{E}$, $\boldsymbol{H}$, $\boldsymbol{B}$ and $\boldsymbol{D}$ are the *electric field*, *magnetic field*, *magnetic flux density vector* and *electric displacement vector*, respectively. $\boldsymbol{J}$ and ρ are the *current density vector* and *charge density*. These six variables are the functions about spatial position $\boldsymbol{r}$ and time t. In addition, there is the law of charge conservation

$$\nabla \cdot \boldsymbol{J} + \frac{\partial \rho}{\partial t} = 0. \tag{3.5}$$

For isotropic medium, the constitutive relations among $\boldsymbol{D}$ and $\boldsymbol{E}$, $\boldsymbol{B}$ and $\boldsymbol{H}$, $\boldsymbol{J}$ and $\boldsymbol{E}$ are given by

$$\boldsymbol{D} = \varepsilon_0 \varepsilon_r \boldsymbol{E}, \qquad \boldsymbol{B} = \mu_0 \mu_r \boldsymbol{H}, \qquad \boldsymbol{J} = \sigma_c \boldsymbol{E}, \tag{3.6}$$

where $\varepsilon_0 \approx 8.854 \times 10^{-12}$ F/m, $\mu_0 = 4\pi \times 10^{-7}$ H/m are the dielectric permittivity and the magnetic permeability of vacuum, and ε_r, μ_r are the relative permittivity and the relative permeability, respectively. σ_c is the electrical conductivity of the medium.

Assuming that an electromagnetic wave propagates in a perfect dielectric medium ($\sigma_c = 0$) and there is no free charge in the medium, which holds for most of the materials for fabricating integrated photonic devices. Eqs. 3.1–3.6 can be simplified as

$$\nabla^2 \boldsymbol{E} = \varepsilon_0 \varepsilon_r \mu_0 \mu_r \frac{\partial^2 \boldsymbol{E}}{\partial t^2}, \tag{3.7}$$

$$\nabla^2 \boldsymbol{H} = \varepsilon_0 \varepsilon_r \mu_0 \mu_r \frac{\partial^2 \boldsymbol{H}}{\partial t^2}. \tag{3.8}$$

The two equations above are the **wave equations** in dielectric medium. Therefore, the propagation speed v of electric field $\boldsymbol{E}$ and magnetic field $\boldsymbol{H}$ in the medium is

$$v = \frac{1}{\sqrt{\varepsilon_0 \varepsilon_r \mu_0 \mu_r}}. \tag{3.9}$$

If the medium is vacuum ($\varepsilon_r = 1, \mu_r = 1$), Eq. 3.9 gives the light speed in vacuum $c = 1/\sqrt{\varepsilon_0 \mu_0}$. Ratio of the two speeds defines the refractive index of the dielectric medium

$$n = \sqrt{\varepsilon_r \mu_r} \approx \sqrt{\varepsilon_r}. \tag{3.10}$$

The approximation holds true as dielectric medium belongs to non-magnetic materials ($\mu_r \approx 1$) generally.

Another important parameter describing the energy transported by the electromagnetic wave, the *Poynting vector*, is defined as

$$\boldsymbol{S} = \boldsymbol{E} \times \boldsymbol{H}. \tag{3.11}$$

Precisely, the vector $\boldsymbol{S}$ represents the energy flow through a unit area in the unit time.

3.1.2. Monochromatic Plane Waves

A monochromatic wave refers to an electromagnetic wave with a constant angular frequency ω (or, a constant wavelength λ), and its field vectors can be characterized by

$$\boldsymbol{E}(\boldsymbol{r}, t) = \mathrm{Re}[\boldsymbol{E}(\boldsymbol{r}) \exp(-\mathrm{i}\omega t)], \tag{3.12}$$

$$\boldsymbol{H}(\boldsymbol{r}, t) = \mathrm{Re}[\boldsymbol{H}(\boldsymbol{r}) \exp(-\mathrm{i}\omega t)]. \tag{3.13}$$

By applying these complex notated fields, the the wave equations (Eqs. 3.7–3.8) can be further written as the Helmholtz equations

$$\nabla^2 \boldsymbol{E} + k^2 \boldsymbol{E} = 0, \tag{3.14}$$

$$\nabla^2 \boldsymbol{H} + k^2 \boldsymbol{H} = 0, \tag{3.15}$$

where the variable k is named as the wave number. Denoting k_0 as the wave number in vacuum, there are the following relations

$$k_0 = \frac{\omega}{c} = \frac{2\pi}{\lambda}, \tag{3.16}$$

$$k = nk_0 = \frac{2n\pi}{\lambda}. \tag{3.17}$$

If the monochromatic wave is also a plane wave and its propagation direction is represented by the wave vector $\boldsymbol{k}$, the complex field amplitude in Eqs. 3.12–3.13 are expressed as

$$\boldsymbol{E}(\boldsymbol{r}) = \boldsymbol{E}_0 \exp(\mathrm{i}\boldsymbol{k}\boldsymbol{r}), \tag{3.18}$$

$$\boldsymbol{H}(\boldsymbol{r}) = \boldsymbol{H}_0 \exp(\mathrm{i}\boldsymbol{k}\boldsymbol{r}). \tag{3.19}$$

Because $\boldsymbol{E}(\boldsymbol{r})$ and $\boldsymbol{H}(\boldsymbol{r})$ also need to satisfy Maxwell's equations, the following relations can be obtained:

$$\boldsymbol{k} \cdot \boldsymbol{E}(\boldsymbol{r}) = 0, \qquad \boldsymbol{k} \cdot \boldsymbol{H}(\boldsymbol{r}) = 0, \tag{3.20}$$

and

$$\boldsymbol{k} \times \boldsymbol{E}(\boldsymbol{r}) = \omega\mu_0 \boldsymbol{H}(\boldsymbol{r}), \qquad \boldsymbol{k} \times \boldsymbol{H}(\boldsymbol{r}) = -\omega\varepsilon_0\varepsilon_r \boldsymbol{E}(\boldsymbol{r}). \tag{3.21}$$

Therefore, it can be inferred that the vectors $\boldsymbol{k}$, $\boldsymbol{E}$ and $\boldsymbol{H}$ are perpendicular to each other. The field vectors $\boldsymbol{E}$ and $\boldsymbol{H}$ locate in a plane which the wave vector is vertical to. Such wave is named as Transverse Electromagnetic (TEM) wave.

3.2. Waveguides

3.2.1. Planar Waveguides

Before investigating the propagation of light in a three-dimensional waveguide, it is necessary to firstly study the propagation mechanism in the planar waveguide depicted in Fig. 3.1. Such waveguide is composed by three different layers, i.e., the core waveguide layer in the middle and the top-cladding and substrate layers at the two sides. Refractive indices of these layers are n_0, n_1 and n_2 from bottom to top, respectively. Usually, these refractive indices have the relation that $n_1 > n_0 \geqslant n_2$. Thicknesses of the top-cladding and substrate layers are much larger than the thickness of the core layer. It is assumed

that the width of the waveguide (dimension in $y-$direction) is also much larger than the thicknesses of all the layers and the waveguide extends to infinity in $x-$direction. Therefore, the electric and magnetic fields of a TEM wave propagating in $z-$direction are functions of x uniquely.

Expressions of the fields can be solved based on the Maxwell's equations and the details can be found in [60, 122]. Several important conclusions from the solutions are summarized below.

Figure 3.1.: Sketch of a planar waveguide composed of three layers with different refractive indices.

The Guided Mode Defining the propagation constant β that represents the wave propagating in $z-$direction, the eigenmode solutions of the Maxwell's equations fall into the following three categories:

1. $0 < \beta < k_0 n_2$: the wave oscillates in all the layers in the form of sinusoidal function and this mode is named as the *radiation mode*.
2. $k_0 n_2 < \beta < k_0 n_0$: the wave oscillates in the core layer and the substrate layer in the form of sinusoidal function whereas experiences exponential attenuation in the top-cladding layer. This mode is named as the *substrate mode*.
3. $k_0 n_0 < \beta < k_0 n_1$: the wave only oscillates in the core layer in the form of sinusoidal function and is exponentially attenuated in the top-cladding and substrate layer. This mode is named as the *guided mode*.

The introductions above explains the reason that the silicon-on-insulator (SOI) wafers are commonly used in the fabrication of integrated photonic circuits. SOI wafers comprise a thin device layer of Si, a box layer of SiO_2 and a thick handle layer of Si from top to bottom. Si has a high refractive index of about 3.5 at the wavelength of 1550 nm [62, 121], whereas SiO_2 has a lower refractive index of about 1.44 at the same wavelength [62]. This large difference of refractive indices perfectly matches the requirements of forming the waveguide when the device layer is chosen as the core layer. The top-cladding layer can be air or another deposited silicon oxide layer.

TE Mode and TM Mode The solutions of the guided mode also gives two types of wave, as drawn in Fig. 3.2. For the first case, the electric field only has a transversal component E_y and the magnetic field has no component in $y-$direction ($H_y = 0$). This type of wave is named as *TE mode*. For the second case, the magnetic field only has a transversal

component H_y while the electric field in this direction is zero. This type of wave is named as *TM mode*.

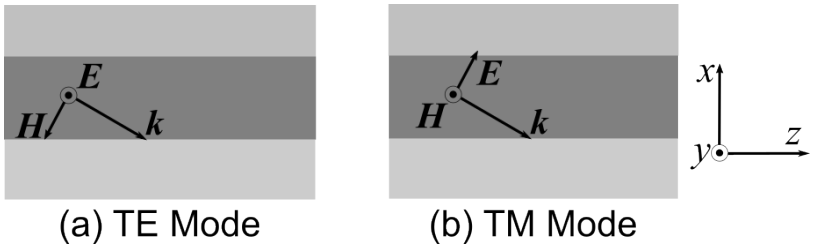


Figure 3.2.: (a) TE and (b) TM propagation in planar waveguide.

Defining the *effective refractive index* of the guided mode as

$$N = \frac{\beta}{k_0}, \tag{3.22}$$

it should satisfy the following eigenvalue equations of planar waveguide [60], which are

for TE mode:

$$(n_1^2 - N^2)^{1/2} k_0 h = m\pi + \arctan\left(\frac{N^2 - n_0^2}{n_1^2 - N^2}\right)^{1/2} + \arctan\left(\frac{N^2 - n_2^2}{n_1^2 - N^2}\right)^{1/2}, \tag{3.23}$$

for TM mode:

$$(n_1^2 - N^2)^{1/2} k_0 h = m\pi + \arctan\left[\frac{n_1^2}{n_0^2}\left(\frac{N^2 - n_0^2}{n_1^2 - N^2}\right)^{1/2}\right] + \arctan\left[\frac{n_1^2}{n_2^2}\left(\frac{N^2 - n_2^2}{n_1^2 - N^2}\right)^{1/2}\right], \tag{3.24}$$

respectively. Here, h is the thickness of the core layer and m is the modal order. Therefore, it is evident that the effective refractive index is not only solely dependent on the material of the core layer, but also influenced by the geometrical dimensions and the mode of the wave.

3.2.2. Channel Waveguides

As discussed in Section 3.2.1, the planar waveguides confine the light propagation only in one direction ($x-$direction in Fig. 3.1). However, such waveguides are not applicable for the actual application in the integrated photonics due to the large footprint. It is preferred that the dimensions in both $x-$ and $y-$directions are comparable, and this type of waveguides is named as the channel waveguides, in which the light is confined in two directions. Usually, the core region that the light is confined in has higher refractive index than its surroundings. Depending on the geometrical shape of the cross-section of the channel waveguide, it can be classified into the rib waveguide, the strip waveguide and

the buried waveguide [60]. There is also a special slot waveguide that the core region has lower refractive index than the surroundings, which has higher fabrication requirements [123]. From the view of the fabrication work in this dissertation, the brief introduction is focused only on the rib and strip waveguides.

The cross-sections of these two types of waveguides are drawn in Fig. 3.3. The light propagates along $z-$direction and is confined in both $x-$ and $y-$directions. Clearly, there is no pure TE or TM mode in neither type of waveguide; the light polarization is a mixture of TEM modes. However, the TEM modes can still be categorized into two types depending on the directions of the major component of the electric field, namely E^x_{mn} and E^y_{mn} modes. The major field components of the former modes are E_x and H_y, which are similar to the fields of TM modes in planar waveguide, therefore these modes are named as *quasi-TM modes*; on the contrary, the major field components of the latter modes are E_y and H_x, and these modes are named as *quasi-TE modes*. The subscripts m and n represent the number of electric field (E_x or E_y) intensity maxima in $x-$ and $y-$directions, respectively. For brevity, the prefix "quasi-" will be omitted in the discussions of these two types of waveguides hereafter.

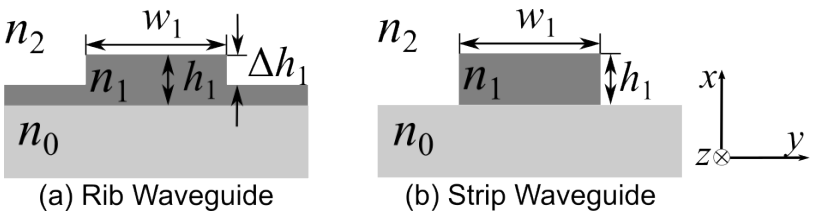


Figure 3.3.: The cross-sections of (a) the rib waveguide and (b) the strip waveguide.

Analytically solving the precise effective refractive index of either rib or strip waveguide is challenging, while the effective index method is an efficient way to calculate the approximated effective refractive index [60]. As an example, the method to calculate the effective refractive index N of the rib waveguide is demonstrated in Fig. 3.4, which contains two equivalent steps: first, the rib waveguide is decomposed into three planar waveguides (Fig. 3.4(a)) and their effective refractive indices, namely N_1 and N_2, are calculated by Eqs. 3.23 and 3.24 for either TE or TM mode. Afterwards, a new planar waveguide with refractive indices N_1, N_2 and N_1 for each layer (Fig. 3.4(b)) is formed and the thickness of the core layer is exactly the width of the rib waveguide. By applying Eqs. 3.23 and 3.24 again, the refractive index of this new planar waveguide can be calculated, which is the refractive index of the rib waveguide.

For the strip waveguide, the equivalent procedures are similar except that considering an additional substrate layer is required to form the first planar waveguide in Fig. 3.4(a), because there is no core layer between the substrate n_0 and top-cladding n_2.

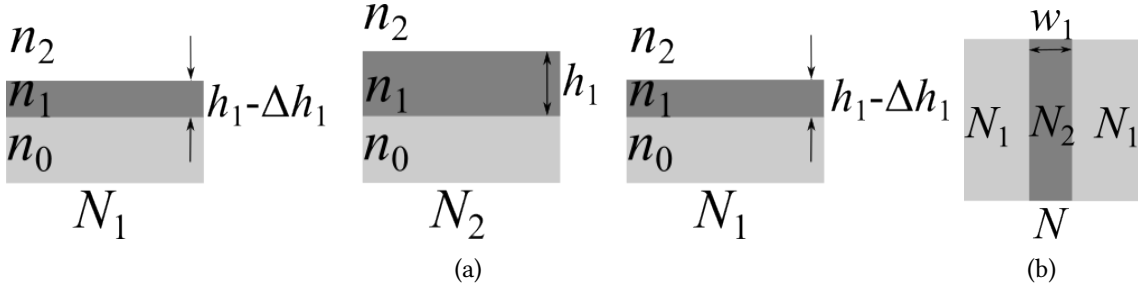


Figure 3.4.: The rib waveguide is equivalent as planar waveguides in two steps to calculate the effective refractive index.

3.2.3. Bent Waveguides

In the introduction above, it is assumed that the light propagates straightly along the $z-$direction inside the waveguides. However, bent waveguides are unavoidable in the design of photonic integrated circuits (PICs) to alter the direction of light propagation. Compared to a straight waveguide, there are more mechanisms that cause losses in bent waveguides and several of them are listed as below [62]:

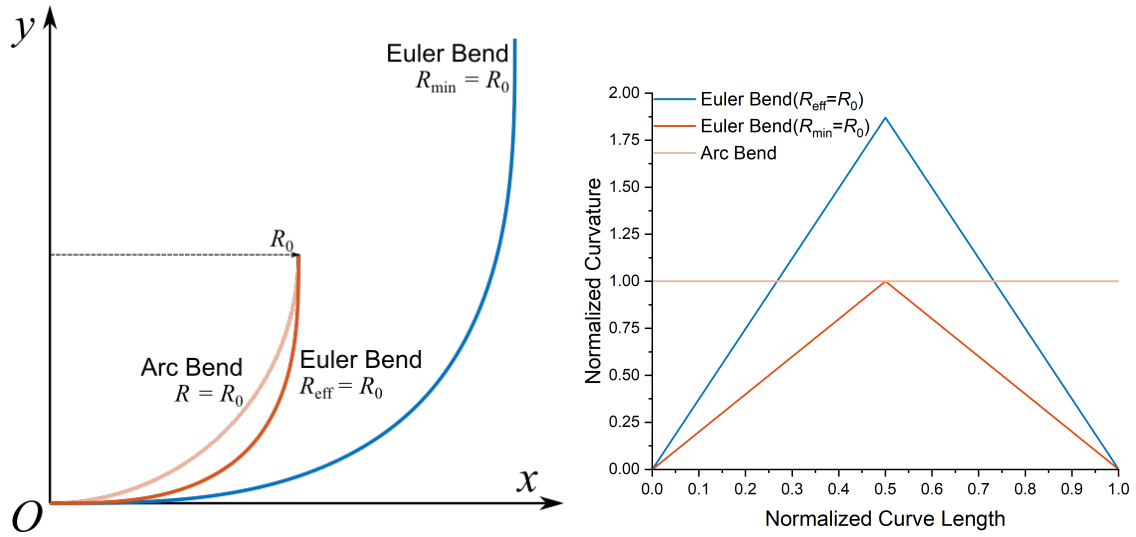
1. **Scattering losses and substrate leakage**: These losses come from fabrication defects, impurities and in material surface roughness, which are small for shorter propagation distance.
2. **Radiative loss**: This loss is closely related to the curvature of bend. Smaller curvature leads to larger radiative loss. Light in TE mode polarization is also beneficial for reducing this loss.
3. **Mode mismatch loss**: This loss comes from imperfect mode overlap when the waveguide curvature changes non-adiabatically and contributes most to the loss in waveguide bends.

The mode mismatch loss is appreciable especially when the straight waveguide is connected by a bent waveguide in arc shape [124]. Although the cross-sections of straight and bent part are the same, there is bending-induced mode mismatch due to the curvature jump at the connecting point. Introducing the Euler bend [124, 125], which possesses an adiabatic radius transition from infinite (straight line) to finite (curved line), is an effective way to reduce the mode mismatch loss significantly. The curvature of the Euler bend $\varphi(s)$ varies linearly with its curve length s:

$$\varphi(s) = \alpha s \tag{3.25}$$

where α is a geometrical parameter. Fig. 3.5(a) intuitively compares the shape of three 90° bends with the same radius R_0, namely an arc bend with fixed curvature along the length ($R = R_0$), an Euler bend[1]with R_{eff} and with R_{min} equaling to R_0. Fig. 3.5(b) demonstrates

relations between the normalized curvature ($\varphi(s) \cdot R_0$) and normalized curve length of these three bends.



(a) Comparison of shapes of bends

(b) Normalized curvatures vs. curve length

Figure 3.5.: Comparison of curve shapes and corresponding normalized curvatures of arc bend ($R = R_0$), Euler bend ($R = R_{\text{eff}}$) and Euler bend ($R = R_{\text{min}}$), respectively.

From Fig. 3.5(a) it can also be learned that the length of the Euler bend is generally larger than the arc bend depending on R_{eff}. Therefore, this introduces additional scattering loss. Besides, the area required for the Euler bend also becomes larger, which is not appealing for a compact design. Some researchers have proposed the *partial Euler bend* design to take advantage of both Euler bend's low loss property and arc bend's compactness [124–126]. Such a bend is constituted by two Euler bends and one arc bend: Transiting from a straight waveguide, it starts with an Euler bend until the curve radius reduces to a specific value R_{min}. An arc bend with R_{min} is then connected to the Euler bend. At the other end of the arc bend, another identical Euler bend is designed so that the curvature can reduce gradually until it becomes zero again, being able to connect to straight waveguide. It has been demonstrated that the total loss can be minimized when the ratio of Euler bend to total bend p is 0.5 [124]. This low-loss bent waveguide shape is also applied in the design of racetrack ring resonators [127]. Detailed geometrical analyses of partial Euler bend can be found in Appendix A.

[1] It is worth noting that the Euler bends described here constitute two identical mathematically defined Euler bends, whose curvature increases monotonically with curve length and shape develops into a spiral. The reason for this difference is that the bent waveguides are designed to connect two straight waveguides in different directions, hence a curvature maximum point is required.

3.3. Directional Coupler

In the design of integrated photonic circuits, there are various fundamental components for different tuning and modulating purposes, which includes the directional coupler, the ring resonators [127], the Mach-Zehnder interferometer and the multi-mode interferometer (MMI) [60, 62]. The introduction here solely focuses on the directional coupler, which is of the research interest of this dissertation in Chapter 7.

A classical directional coupler consists of two identical waveguides that have the same propagation constants and are aligned in parallel. Due to the oscillation features of the electromagnetic waves, there is coupling effect between the light waves in the two separated waveguides when the gap between them is close enough. A demonstrative directional coupler is drawn in Fig. 3.6, which comprises a bus waveguide and a receiving waveguide. The gap distance between them and the coupling length are denoted by g and L_c. Assuming a wave is input from the bus waveguide, after propagating a certain distance, the energy of the wave is redistributed in the two waveguides due to coupling effect and be output through the two output ports. In case of perfect coupling, the wave is completely transferred to the receiving waveguide and there is no output from the bus waveguide.

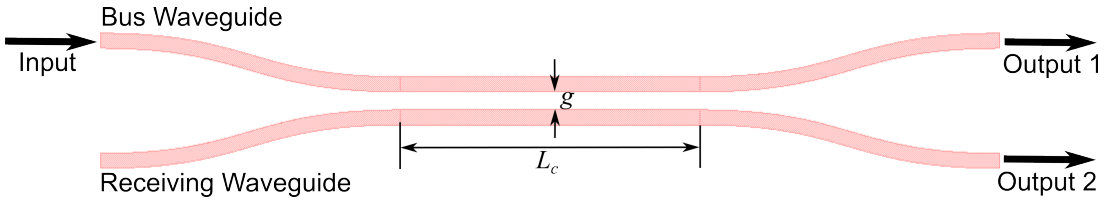


Figure 3.6.: Sketch of a directional coupler.

Based on the "supermode" method [62], the eigenmodes of directional coupler consist of a symmetric mode β_s and an asymmetric mode β_a. The different phase velocities of the two modes leads to a phase difference $\Delta\varphi$ between the modes. When $\Delta\varphi = 2n\pi \quad (n = 0, 1, 2, \ldots)$, the wave power is concentrated in the bus waveguide; while if $\Delta\varphi = (2n + 1)\pi \quad (n = 0, 1, 2, \ldots)$, the wave power in the receiving waveguide reaches maximum. Therefore, the coupling length for a full power transfer is

$$L_c = \frac{\lambda}{2(n_s - n_a)}, \tag{3.26}$$

where n_s and n_a are the effective refractive indices of the symmetric mode and the asymmetric mode, respectively.

The coupling efficiency which denotes the proportion of power coupling into the receiving waveguide is

$$\kappa^2 = \frac{P_{\text{coupled}}}{P_{\text{total}}} = \sin^2\left(\frac{\pi}{2} \cdot \frac{z}{L_c}\right), \tag{3.27}$$

where z is the distance that the light has traveled in the directional coupler.

There are also different variants based on the classical design to improve the performance of coupling. For example, an adiabatic design of directional coupler has been investigated recently to achieve wider operation bandwidth [128, 129]. These structures are beyond the discussions of this dissertation.

3.4. Grating Coupler

One of the most important components of integrated photonic circuits that has not been introduced yet is the optical input/output (I/O) port, which can guide the light from the external optical fibers to the on-chip waveguides and reversely. The challenge comes from that there is enormous mode mismatch between the optical fiber and the waveguides, which originates from the large difference of the feature sizes. The typical diameter of an optical fiber is around 10 μm while the dimensions of the cross-section of a typical waveguide are 500 nm × 220 nm. Hence, it is necessary to have a component adjusting the modes between the waveguide and the optical fiber.

There have been different approaches proposed to tackle this problem [121], and two of the most commonly adopted methods are by designing the grating couplers and the edge couplers. Having the advantages of simple fabrication processes and flexibility of position [62, 130], grating couplers are selected for designing the photonic circuit of the work of this dissertation.

Apart from some specially designed apodized grating couplers [131–134], the general grating couplers are periodically etched structures on the device layer of SOI wafers that can diffract light. In brief, the working mechanism of grating couplers can be described by the Bragg condition [135] as

$$\beta_{\mathrm{wg}} - k_z = \frac{2m\pi}{p}, \qquad m = 1, 2, 3, \ldots, \tag{3.28}$$

where β_{wg} is the propagation constant of the waveguide, k_z is the projected component of the incident wave vector $\boldsymbol{k}$ in the z−direction, m is diffraction modal order and p is the period of the grating. These variables are demonstrated in the schematic diagram Fig. 3.7.

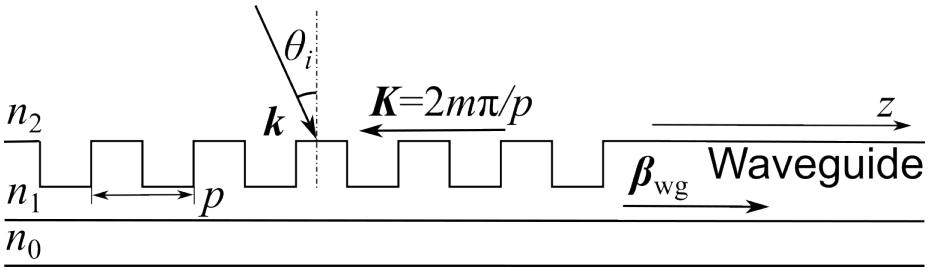


Figure 3.7.: Illustrative Bragg condition for periodic grating coupler.

Substituting with Eqs. 3.17 and 3.22 and taking the first mode of diffraction, the Eq. 3.28 becomes

$$N - n_2 \sin\theta_i = \frac{\lambda}{p}, \tag{3.29}$$

where N is the effective refractive index of the waveguide and θ_i is the coupling angle between the incident wave and the vertical direction (if the light propagates horizontally in the waveguide). Evidently, the grating coupler is a polarization-dependent structure and is sensitive to the wavelength of the light. If the top-cladding is air, the coupling angle is expressed as

$$\sin\theta_i = N - \lambda/p, \tag{3.30}$$

which is a helpful equation to determine the coupling angle in transmission tests.

3.5. Summary

This chapter has given an overview of the basics of the integrated photonics from the fundamental Maxwell's equations to some on-chip photonic components that are used in the work of this dissertation, which include the rib and strip waveguides, the directional coupler and the grating coupler. The concepts, definitions and equations introduced form the basis for the design and characterization work of the photonic circuits in Chapter 7.

4. Mechanical and Thermal Analyses of Multilayer Beam Structures

In Si-based MEMS and NEMS technologies, freely movable and mechanically active "actuator" structures are usually based on beam structures with at least two layers, i.e., Si layer as the common substrate and a layer of a transducer material that converts input energy into work. Therefore, it is of importance to understand the mechanical performance and temperature distribution of multilayer beam structures, especially when heating is involved in the actuation mechanism. Although having the shortcomings of limitation to rather simple material model and structural geometry, building an analytical model is still a simple and fast method useful for an overall understanding on the influences of all involved parameters. In this chapter, the bending behavior of multilayer beams due to inelastic strains is firstly analyzed based on Euler-Bernoulli beam theory. Next, the temperature distribution along such a structure is obtained by solving the equation of Fourier's law, and this analytical model is further validated by experiments at macroscale. A wing structure design proposed recently is used to overcome the large temperature gradient along beams at the nanoscale. The mechanical influence of introducing this design is studied in the following section based on the energy method.

4.1. Mechanical Analysis

4.1.1. Stress, Strain Distribution and Deflection

Considering a general beam illustrated in Fig. 4.1, the elastic modulus $E(y)$ and beam width $b(y)$ vary along the thickness direction. According to the assumption of Euler-Bernoulli beam theory, the cross-section of the beam keeps flat and remains vertical to the beam axis after bending. The stress distribution along the thickness direction can be therefore expressed linearly:

$$\sigma(y) = E(y)[\varepsilon_0 + \varphi y - \varepsilon_{\text{inelastic}}(y)], \tag{4.1}$$

where ε_0 is the strain at the centroid z_c in the cross-section of the beam, φ is the curvature of the bent beam and $\varepsilon_{\text{inelastic}}(y)$ is the inelastic strain, which includes strains due to the thermal expansion effect and the residual stress.

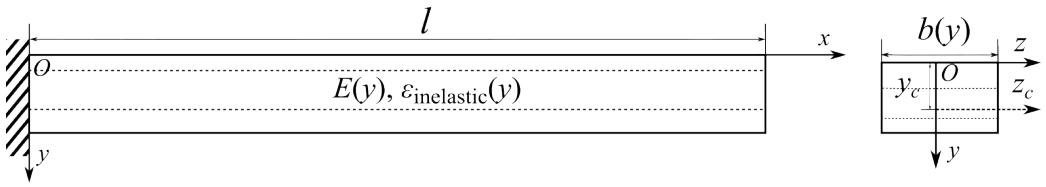


Figure 4.1.: Sketch of a general Euler-Bernoulli beam model. Elastic modulus, inelastic strains and width vary in thickness direction. The distance between the top surface of beam and centroid in the cross section is marked as y_c.

In Fig. 4.1, the $z-$axis is placed at the top surface of the beam preliminarily. In the following analysis, the $z-$axis is shifted to the z_c-axis crossing the centroid in the cross-section. The distance y_c between the top surface and centroid is

$$y_c = \frac{\int E(y) y b(y)\,\mathrm{d}y}{\int E(y) b(y)\,\mathrm{d}y}. \tag{4.2}$$

In the case of a cantilever without external load on the beam in a quasi-static deformation, the boundary conditions are that both external force and moment equal to zero

$$\begin{aligned} F_{\text{external}} = 0: &\qquad \int \sigma(y) b(y)\,\mathrm{d}y = 0; \\ M_{\text{external}} = 0: &\qquad \int y\sigma(y) b(y)\,\mathrm{d}y = 0. \end{aligned} \tag{4.3}$$

Combining Eqs. 4.1–4.3 gives ε_0 and φ as

$$\varepsilon_0 = \frac{\int E(y)\varepsilon_{\text{inelastic}}(y) b(y)\,\mathrm{d}y}{\int E(y) b(y)\,\mathrm{d}y}, \tag{4.4}$$

$$\varphi = \frac{\int E(y)\varepsilon_{\text{inelastic}}(y) y b(y)\,\mathrm{d}y}{\int E(y) y^2 b(y)\,\mathrm{d}y}. \tag{4.5}$$

The deflection value at the free-end of a cantilever-shaped beam actuator is usually of great interest, which is dependent on the beam curvature and length. As depicted in Fig. 4.2, the deformed shape of the beam can be defined by variables (x, y, ϕ) in a global planar Cartesian coordinate system Oxy, where ϕ is the rotating angle of cross-section at point (x, y) and also equals to the angle between $x-$axis and tangent line of deformed beam at this point. The parametric equations of these three variables based on the arc distance s from free-end of cantilever to the point (x, y) along the beam and the curvature φ therefore are

$$\frac{\mathrm{d}x(s)}{\mathrm{d}s} = -\cos\phi(s); \quad \frac{\mathrm{d}y(s)}{\mathrm{d}s} = -\sin\phi(s); \quad \varphi = -\frac{\mathrm{d}\phi(s)}{\mathrm{d}s}, \tag{4.6}$$

and the corresponding boundary conditions are

$$x(l) = 0; \quad y(l) = 0; \quad \phi(l) = 0. \tag{4.7}$$

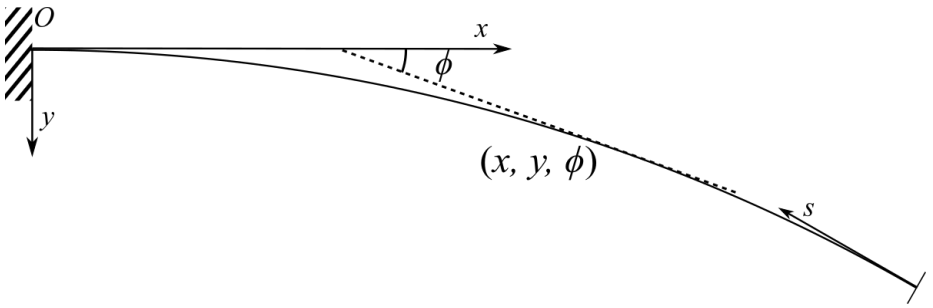


Figure 4.2.: Sketch of a deformed beam. Variables (x, y, ϕ) are introduced to define the flexure curve of the beam.

Assuming that the curvature φ is independent of the position on the beam and the elongation of the beam is negligible, from Eqs. 4.6–4.7 the deformed beam shape is given as

$$\begin{aligned} x(s) &= \frac{1}{\varphi} \sin\left[\varphi(l-s)\right], \\ y(s) &= \frac{1}{\varphi}(1 - \cos\left[\varphi(l-s)\right]). \end{aligned} \tag{4.8}$$

The deflection δ at the free-end of the cantilever is

$$\delta = y(0) = \frac{1}{\varphi}[1 - \cos(\varphi l)]. \tag{4.9}$$

It is worth noting that the assumption of small deflection is not considered in the analysis above, therefore Eqs. 4.1–4.9 remain valid in the case of large deflection. However, the assumption of constant curvature does not hold under certain circumstances. For example, there is a temperature gradient in Joule-heated cantilever actuators along the beam length direction, which will be demonstrated in Section 4.2. Therefore, $\varepsilon_{\text{inelastic}}(y)$ induced by different thermal expansion coefficients effect becomes dependent on the position in the length direction. The external moment leads to a constant curvature, whereas the external force results in a position-dependent curvature as well. Solutions of Eq. 4.6 in such a scenario are based on functions of elliptic integral, which have been investigated in previous work [136–138], and are not further discussed here.

4.1.2. Equivalent Treatment

To further simplify the analyses of multilayer beam structures, the following equivalent parameters are taken based on the parameters from all the layers.

Geometrical and Material Parameters These equivalent parameters are introduced so that the multilayer structure can be simplified as a single-layer structure of one material, which are

beam height:

$$h_{\mathrm{eq}} = \int \mathrm{d}y, \tag{4.10}$$

beam width:

$$b_{\mathrm{eq}} = \frac{\int b(y)\,\mathrm{d}y}{\int \mathrm{d}y}, \tag{4.11}$$

elastic modulus:

$$E_{\mathrm{eq}} = \frac{\int E(y)b(y)\,\mathrm{d}y}{\int b(y)\,\mathrm{d}y}, \tag{4.12}$$

bending stiffness:

$$(EI)_{\mathrm{eq}} = \int E(y)y^2 b(y)\,\mathrm{d}y, \tag{4.13}$$

respectively.

External Loads The multilayer beam structure described in Fig. 4.1 are bent because of the effect of inelastic strain, which can be replaced by an equivalent external force and moment applied on the free-end of the beam resulting in the same deformation shape, i.e., the tensile strain at the centroid ε_0 and the curvature of the beam φ remain identical. Therefore, the equivalent force is

$$F_{\mathrm{eq}} = \int E(y)\varepsilon_0 b(y)\,\mathrm{d}y, \tag{4.14}$$

combining with Eq. 4.4, it equals to

$$F_{\mathrm{eq}} = \int E(y)\varepsilon_{\mathrm{inelastic}}(y)b(y)\,\mathrm{d}y. \tag{4.15}$$

Because the bending equation of Euler-Bernoulli beam is [139]

$$\varphi = \frac{M}{EI}, \tag{4.16}$$

combining with Eq. 4.5 and Eq. 4.13, the equivalent moment is

$$M_{\mathrm{eq}} = \int E(y)\varepsilon_{\mathrm{inelastic}}(y)yb(y)\,\mathrm{d}y. \tag{4.17}$$

It should be highlighted that Eq. 4.16 implies the assumption of small deflection.

4.1.3. Case Study: Bimorph Cantilever Beam and Stress Measurement

Giving a bimorph cantilever beam similar to the beam shown in Fig. 4.1, its width, length and layer thicknesses are b, l, h_1 and h_2, respectively. The elastic moduli and inelastic strains of the two layers are $E_1, E_2, \varepsilon_1^{\text{in}}, \varepsilon_2^{\text{in}}$. Therefore, the position of the centroid of the cross-section defined by Eq. 4.2 is

$$y_c = \frac{E_1 h_1^2 + E_2(2h_1 h_2 + h_2^2)}{2(E_1 h_1 + E_2 h_2)}. \tag{4.18}$$

The curvature of the cantilever beam due to inelastic strain is then obtained by applying Eq. 4.5 as

$$\varphi = \frac{6(h_1 + h_2)(\varepsilon_2^{\text{in}} - \varepsilon_1^{\text{in}})}{h_2^2 K}, \tag{4.19}$$

where

$$K = 4 + 6\frac{h_1}{h_2} + 4\left(\frac{h_1}{h_2}\right)^2 + \frac{E_1}{E_2}\left(\frac{h_1}{h_2}\right)^3 + \frac{E_2}{E_1}\frac{h_2}{h_1}.$$

Eq. 4.19 acts as the basis of stress measurement in thin-film engineering, as the residual stress or strain in film can be explicitly revealed by the change of curvature of beam; by combining with Eq. 4.9, it can be further demonstrated by the change of deflection [140–143].

Additionally, in case that the film is deposited on a wafer scale, i.e., the two in-plane dimensions are much larger than dimension in out-of-plane direction, the bilayer system is in biaxial stress state, and the elastic modulus needs to be modified accordingly due to the effect of Poisson's ratio as

$$E' = \frac{E}{1 - \mu}, \tag{4.20}$$

where E' is the biaxial elastic modulus and μ is Poisson's ratio. Assuming that there is no residual strain in the substrate ($\varepsilon_2^{\text{in}} = 0$) and its thickness is much larger than the film ($h_2 \gg h_1$), based on Eq. 4.19 and Eq. 4.20, the relation between the residual stress in deposited thin-film σ_f and the curvature change $\Delta\varphi$ after deposition is expressed as

$$\sigma_f = -\frac{E_2 h_2^2}{6(1 - \mu_2)h_1}\Delta\varphi, \tag{4.21}$$

which is exactly the famous Stoney equation [86, 140–142]. The minus sign here indicates the direction of the curvature: a positive curvature corresponds to a concave shape indicating compressive residual stress in the film, whereas a negative curvature corresponds to a convex shape indicating tensile stress in the film.

4.2. Thermal Analysis

4.2.1. Analytical Model

The following presumptions are made beforehand when considering this model:

- Ambient air is at isothermal state denoted by t_∞.
- The substrate shares the same temperature as the ambient air.

The differential form of Fourier's law in cylindrical coordinate system (r, φ, z)[144] is written as

$$\rho c \frac{\partial t}{\partial \tau} = \frac{1}{r}\frac{\partial}{\partial r}(\lambda r \frac{\partial t}{\partial r}) + \frac{1}{r^2}\frac{\partial}{\partial \varphi}(\lambda \frac{\partial t}{\partial \varphi}) + \frac{1}{z}\frac{\partial}{\partial z}(\lambda \frac{\partial t}{\partial z}) + \dot{\Phi}, \tag{4.22}$$

where

- t is temperature,
- ρ is density,
- c is specific heat capacity,
- τ is time,
- λ is thermal conductivity,
- $\dot{\Phi}$ is net heat generated by inner heat source in unit volume and unit time.

Because the length of the beam is usually much larger than its width and thickness, it can be assumed that the temperature gradient in the plane of the beam's cross-section is negligible. If thermal conductivity λ is independent of the position and in a steady situation (time-independent), Eq. 4.22 can be simplified as

$$\lambda \frac{\mathrm{d}^2 t}{\mathrm{d}z^2} + \dot{\Phi} = 0. \tag{4.23}$$

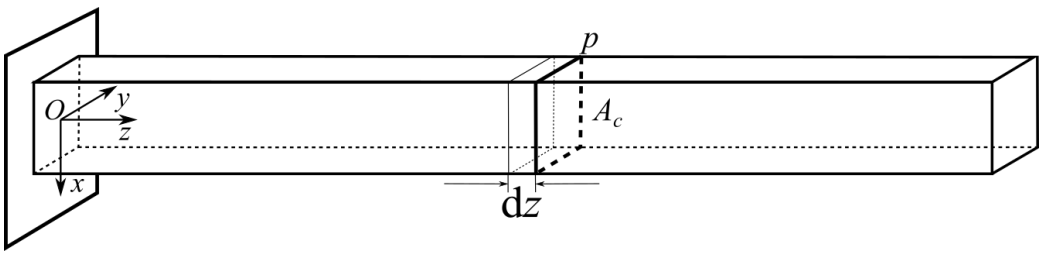


Figure 4.3.: Sketch of 1-D thermal model of cantilever beam

For this 1-D situation, considering a finite element section dz shown in Fig. 4.3, the net heat generated in dz by Joule heating and heat dissipated to surrounding environment in unit time, which includes heat convection and radiation to the air, is

$$\Phi = -h(t - t_\infty)p\,\mathrm{d}z + I^2 \rho_{\mathrm{eff}} \frac{\mathrm{d}z}{A_c}. \tag{4.24}$$

Therefore $\dot{\Phi}$ in unit volume is

$$\dot{\Phi} = -h(t - t_\infty)p + I^2 \frac{\rho_{\text{eff}}}{A_c}. \tag{4.25}$$

The meaning of the new notations here is:

- h is the combined heat convection and radiation transfer coefficient,
- p is the perimeter of the beam's cross-section,
- I is the current through the beam,
- ρ_{eff} is the effective resistivity and it can change along $x-$direction,
- A_c is the area of the cross-section.

The Dirichlet boundary condition at the fixed end of the beam and the Neumann boundary condition at the free end are introduced respectively:

$$\begin{aligned} z = 0: &\qquad t = t_\infty; \\ z = l: &\qquad \frac{\mathrm{d}t}{\mathrm{d}z} = 0. \end{aligned} \tag{4.26}$$

Introducing the concept of *excess temperature* as

$$\theta = t - t_\infty, \tag{4.27}$$

and combining Eqs. 4.23–4.26, the temperature profile in beam structure is solved as

$$\theta(z) = \frac{n^2}{m^2} \left\{ 1 - \frac{\cosh[m(z-l)]}{\cosh(ml)} \right\}, \tag{4.28}$$

and the corresponding temperature gradient is

$$\frac{\mathrm{d}\theta}{\mathrm{d}z} = -\frac{n^2}{m} \frac{\sinh[m(z-l)]}{\cosh(ml)}, \tag{4.29}$$

where

$$\begin{aligned} m &= \sqrt{\frac{ph}{\lambda A_c}} \\ n &= \sqrt{\frac{I^2 \rho_{\text{eff}}}{\lambda A_c^2}}. \end{aligned} \tag{4.30}$$

The maximum temperature occurs at the free end ($z = l$) of the beam and the value is

$$\theta_{\max} = \frac{n^2}{m^2} \left[1 - \frac{1}{\cosh(ml)} \right]. \tag{4.31}$$

It is crucial to highlight that the model imposes no restrictions on the shape of cross-section, unlike the mechanical model in Section 4.1. Cross-sections with the same perimeter and area leads to the same temperature profile.

Quantifying the energy dissipation mechanism is also of large interest in design of micro- and nanoactuators. As introduced above, the heat generated by electric current is dissipated to the substrate by conduction and to the ambient air by convection and radiation, which are denoted as $P_{\text{substrate}}$ and P_{air} respectively. They can be calculated based on Fourier's law and Newton's law of cooling [144]

$$P_{\text{substrate}} = \frac{n^2}{m}\lambda A_c \tanh(ml), \tag{4.32}$$

$$P_{\text{air}} = \frac{n^2}{m^2} phl \left[1 - \frac{\tanh(ml)}{ml}\right]. \tag{4.33}$$

In a steady state, the total dissipated energy, P_{total}, equals to the generated electric energy

$$P_{\text{total}} = \frac{\rho l I^2}{A_c}, \tag{4.34}$$

according to the law of conservation of energy. Hence, the ratio of dissipated heat through the substrate is

$$\frac{P_{\text{substrate}}}{P_{\text{total}}} = \frac{\tanh(ml)}{ml}. \tag{4.35}$$

To validate the temperature profile from Eq. 4.28, a 3D-finite element model of Joule-heated Cu/Si bimorph double-beam cantilever is built in Comsol Multiphysics® as depicted in Fig. 4.4. The length, width and thicknesses of the Si and Cu layers in each beam are 10 μm, 400 nm, 220 nm and 300 nm, respectively. A voltage is applied between two fixed ends of the double-beam cantilever. The extracted temperature profile at the center of the cross-section along the direction of beam length shows excellent agreement with the calculated results, indicating that this analytical model is precise enough to analyze this Joule-heating scenario.

Temperature profiles of various beam lengths from 2.5 μm to 30 μm are also plotted in the same figure. It is interesting to notice that the maximum temperature at the free-ends of the beams remains almost unchanged in such a large length range. This shows that the beam length can barely lower the temperature gradient in Joule-heated micro- and nanoactuators. The corresponding power ratio from Eq. 4.35 drops slightly from nearly 100% to 99.6% in this length range, indicating that heat conduction to the substrate plays the major role in the heat dissipation mechanism.

A similar study of varying width from 0.15 μm to 5 μm is also carried out based on Eqs. 4.31 and 4.35, and there is no noticeable change of neither the corresponding $\theta_{\max}$ nor power ratio, which further confirms that variations of dimensions have a minor influence on the large temperature difference and the heat dissipation mechanism of Joule-heated micro- and nanoactuators.

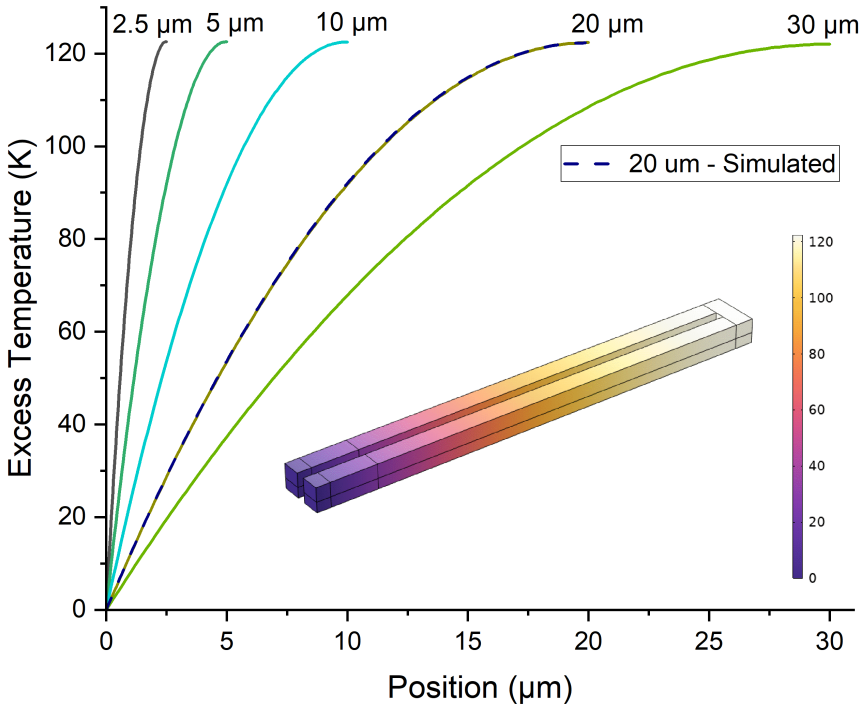


Figure 4.4.: Calculated temperature profiles of cantilevers with different beam length. Joule-heated 20 µm-cantilever is simulated with result plotted in dashed line for comparison.

4.2.2. Experimental Validation

Due to the resolution limitation of most of the existing temperature measurement techniques, it is challenging to in-situ measure the real-time temperature distribution of micro- and nanoactuators. Therefore, a Joule-heated structure at macroscale is built and its temperature distribution is captured by an infrared camera, which is then compared with the analytical model. [1]

As shown in Fig. 4.5(a), a tin wire of 75 mm in length is structured into a double-beam cantilever shape. The diameter of the wire is 0.5 mm. The two ends of the wire are fixed on the steel pads that are large enough to be used as the heat sink. A current is applied through the wire to Joule-heat the structure and the temperature distribution along the beam is captured by the infrared camera FLIR A665sc-25, which has a detector pitch of 17 µm. Fig. 4.5(b) demonstrates the temperature distribution of a part of the beam when being Joule-heated. It can be noticed that the temperature gradient in the diameter direction is negligible, which validates the one-dimensional assumption of the analytical model.

The experimental and analytical temperature distribution curves of this double-beam cantilever structure are depicted in Fig. 4.6, and there is a noticeable gap between the two curves that is irrelevant to the position. This difference can be explained by the failure of the Dirichlet boundary condition in Eq. 4.26, which assumes that there is no temperature

[1] Experiments conducted by master's student Nengqi Zhang.

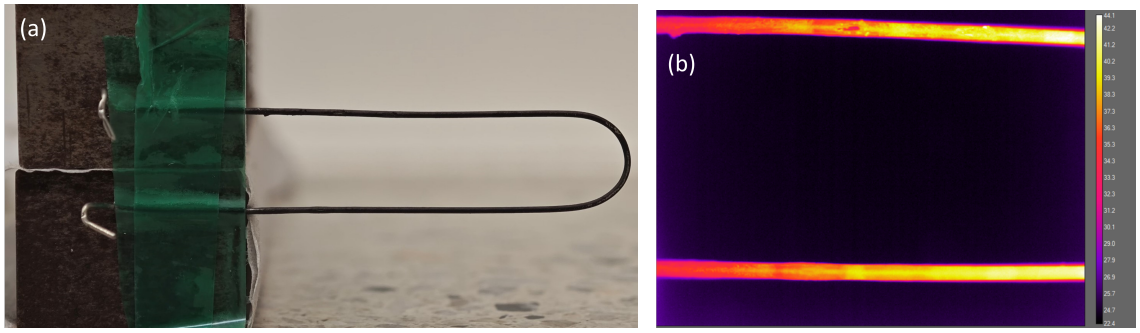


Figure 4.5.: Experimental validation of the thermal model by a Joule-heated double-beam cantilever. (a) A tin wire is folded into the shape of a double-beam cantilever structure. Its left ends are connected to the steel pads as the heat sink. (b) The wire is Joule-heated and the temperature distribution is captured by the infrared camera.

gradient between the fixed end of the beam and the substrate at environment temperature. However, due to the non-ideal heat conduction between the fixed ends of wire and the steel pads, the former are at about 33 °C while the latter are at room temperature of 24 °C. The corrected analytical temperature distribution is also plotted in Fig. 4.6, which is slightly higher than the experimental results. This error can be attributed to the inaccuracy of the heat convection coefficient adopted in the analytical model.

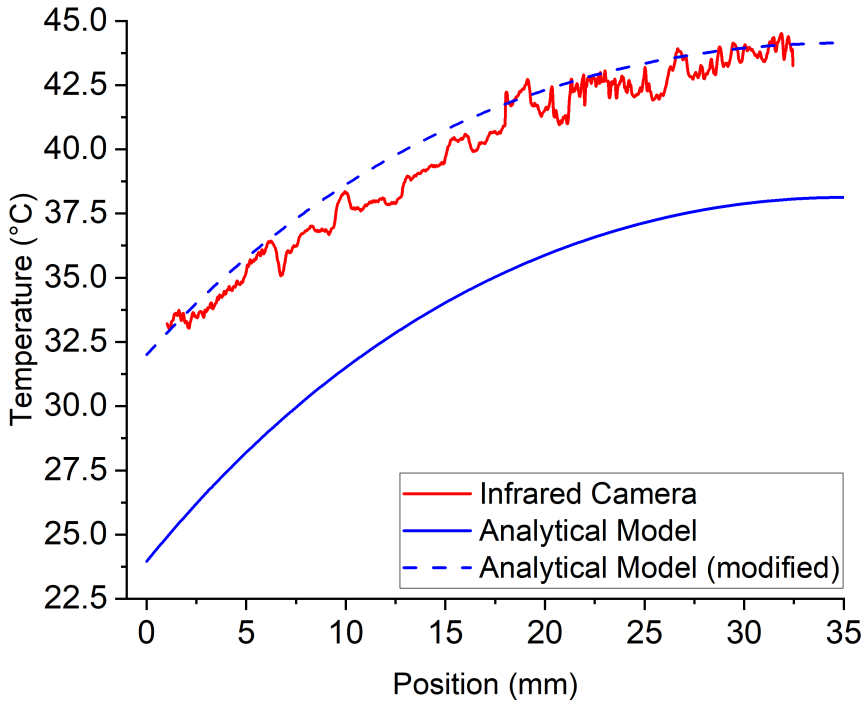
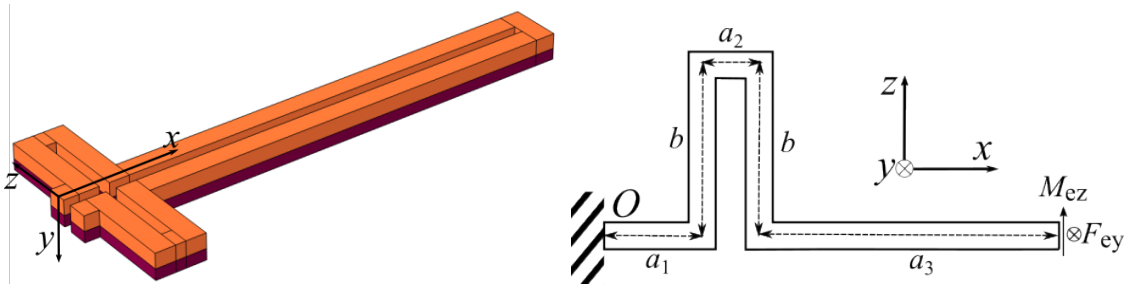


Figure 4.6.: Experimental and analytical temperature distribution of the tin wire double-beam cantilever structure when being Joule-heated. The analytical result is modified by correcting the boundary condition.

4.3. Wing Structure Design

In Section 4.2, it has been revealed that there is significant temperature gradient in Joule-heated beam cantilever structures along length direction, which is unfavorable for the full performance of corresponding SMA/Si beam actuators as the temperature of only a part along the beam length can be above the A_f temperature of phase transformation.

A wing structure as depicted in Fig. 4.7 has been proposed to tackle this problem [145, 146]. The idea behind such design is to fold the undesirable low-temperature regime on the beam to a wing shape. It has been demonstrated that this design is beneficial to reduce power consumption of bimorph Si/SMA actuator up to 48%, compared to conventional double-beam cantilever design [146]. However, the influence of the wing-shape design on the mechanical performance of the SMA/Si beam actuator has not been investigated. In the following, the influence on structural bending stiffness will be presented based on the energy method.



(a) 3D model (b) 2D model with external loads

Figure 4.7.: (a) 3D and (b) 2D model of a multilayer double-beam cantilever with wing structure design. Due to symmetry, only half of the structure is considered. The structural dimensions and external loads are marked out.

For a 1D deformable beam-shaped structure of linear elastic materials under different types of external loads, as depicted in Fig. 4.8, strain energies from tensile, torsional, bending and shearing deformations are expressed as

$$V_{\text{tensile}} = \int_l \frac{F_N^2(x)\,\mathrm{d}x}{2EA_c}, \tag{4.36}$$

$$V_{\text{torsion}} = \int_l \frac{T^2(x)\,\mathrm{d}x}{2GJ}, \tag{4.37}$$

$$V_{\text{bending}} = \int_l \frac{M^2(x)\,\mathrm{d}x}{2EI}, \tag{4.38}$$

$$V_{\text{shearing}} = \int_l \frac{kF_s^2(x)\,\mathrm{d}x}{2GA_c}, \tag{4.39}$$

respectively. The meanings of notations used above are explained as below:

$F_N(x), T(x), M(x)$ **and** $F_s(x)$ Tensile force, torsional moment, bending moment and shearing force,

E **and** G Elastic and shear modulus of material. In the case of multilayer structure, they represent equivalent elastic and shear modulus,

A_c, I **and** J Area, area moment of inertia and torsion constant of the cross-section of beam,

k A factor related to the shape of cross-section without unit. For a beam with rectangular cross-section, this factor equals to 6/5.

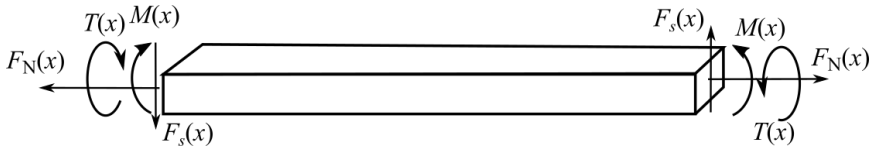


Figure 4.8.: 1D beam-shape structure with different external loads: tensile force $F_N(x)$, torsional moment $T(x)$, bending moment $M(x)$ and shear force $F_s(x)$.

Given a structure comprised by n parts of beams, the total structural strain energy is calculated by summing up energies of all beams

$$U = \sum_i \left(V^i_{\text{tensile}} + V^i_{\text{torsion}} + V^i_{\text{bending}} + V^i_{\text{shearing}} \right). \tag{4.40}$$

Castigliano's second theorem states that, for linearly elastic structure of random shape, the partial derivative of strain energy with respect to generalized force F_i gives the generalized displacement δ_i in the direction of F_i, which is expressed as:

$$\delta_i = \frac{\partial U(F_1, F_2, ..., F_i, ...)}{\partial F_i}. \tag{4.41}$$

Therefore, the corresponding structural stiffness is

$$k_i = \frac{F_i}{\delta_i}. \tag{4.42}$$

Considering a cantilever beam with wing structure shown in Fig. 4.7(b), note that only half of the structure in Fig. 4.7(a) is analyzed due to symmetry, the structure undertakes external loads at its free-end, i.e., the force F_{ez} in $y-$direction and the moment M_{ez} in $z-$direction, and it is fixed at the other end. The lengths of the beam are a_1, a_2, a_3 in length direction and wing length is b, respectively. Denoting that the beam length $l = \sum a_i$ and the ratio of wing length to beam length $\eta = b/l$. The following assumptions are made:

- The lengths of all parts, i.e., a_1, a_2, a_3 and b, are much larger than the beam width w. Hence, the connecting points of different parts can be seen as a point.
- The material properties (E, G) and geometrical parameters (A, I, J) are independent of the position in the length direction of the beam.

Case I: Applying an external moment M_{ez} The external moment arouses rotation of the cross-section at the free-end of the beam, the rotated angle $\phi_{\text{wing}}^{M_{ez}}$ is therefore the generalized displacement corresponding to M_{ez}. Combining Eqs. 4.36–4.42 gives the rotating stiffness as

$$k_{\text{wing}}^{M_{ez}} = \left(\frac{l}{EI} + \frac{2b}{GJ}\right)^{-1}. \tag{4.43}$$

When there is no wing structure in the beam, i.e., $b = 0$, the corresponding stiffness of the cantilever beam is

$$k_{\text{straight}}^{M_{ez}} = \left(\frac{l}{EI}\right)^{-1}, \tag{4.44}$$

and the ratio $\gamma^{M_{ez}}$ between these two structures is

$$\gamma^{M_{ez}} = \frac{k_{\text{wing}}^{M_{ez}}}{k_{\text{straight}}^{M_{ez}}} = \left(1 + 2\eta\frac{EI}{GJ}\right)^{-1}. \tag{4.45}$$

This external moment also induces deflection $\Delta_{\text{wing}}^{M_{ez}}$ at the free-end, which can be obtained based on the principle of virtual work

$$\Delta_{\text{wing}}^{M_{ez}} = \frac{M_{ez}l^2}{2EI} + \frac{M_{ez}}{GJ}(a_2 + 2a_3)b. \tag{4.46}$$

Likewise, deflection of conventional cantilever beam $\Delta_{\text{straight}}^{M_{ez}}$ can also be got by setting $b = 0$

$$\Delta_{\text{straight}}^{M_{ez}} = \frac{M_{ez}l^2}{2EI}, \tag{4.47}$$

the ratio $\zeta^{M_{ez}}$ between them is

$$\zeta^{M_{ez}} = \frac{\Delta_{\text{wing}}^{M_{ez}}}{\Delta_{\text{straight}}^{M_{ez}}} = 1 + 2\eta\frac{a_2 + 2a_3}{l}\frac{EI}{GJ}. \tag{4.48}$$

For a monomorph beam and assuming that $a_3 \gg a_1, a_2$, Eq. 4.45 and Eq. 4.48 are simplified as

$$\gamma^{M_{ez}} \approx \left[1 + \eta\frac{1+\mu}{3\beta}\left(\frac{h}{w}\right)^2\right]^{-1}, \tag{4.49}$$

and

$$\zeta^{M_{ez}} \approx 1 + \eta\frac{2(1+\mu)}{3\beta}\left(\frac{h}{w}\right)^2, \tag{4.50}$$

respectively. Here, μ is Poisson's ratio, h and w are the height and width of the cross section of the beam, β is a geometrical parameter related to the torsion of beam with rectangular cross-section [147].

Case II: Applying an external force F_{ey} The external force arouses deflection at the free-end of the beam, the deflected value $\Delta_{\text{wing}}^{F_{ey}}$ is therefore the generalized displacement corresponding to F_{ey}. Similarly, the deflecting stiffnesses of the beam with and without wing structure are

$$k_{\text{wing}}^{F_{ey}} = \left\{ \frac{l^3}{3EI} + \frac{2b^3}{3EI} + \frac{a_2 b^2}{GJ} + \frac{[(a_2+a_3)^2 + a_3^2]b}{GJ} + \frac{k(l+2b)}{GA_c} \right\}^{-1}, \tag{4.51}$$

and

$$k_{\text{straight}}^{F_{ey}} = \left(\frac{l^3}{3EI} \right)^{-1}, \tag{4.52}$$

respectively. The stiffness ratio $\gamma^{F_{ey}}$ is

$$\gamma^{F_{ey}} = \frac{k_{\text{wing}}^{F_{ey}}}{k_{\text{straight}}^{F_{ey}}} = \left[1 + 2\eta^3 + 3\eta^2 \frac{a_2}{l} \frac{EI}{GJ} + 3\eta \frac{(a_2+a_3)^2 + a_3^2}{l^2} \frac{EI}{GJ} + 3k \frac{1+2\eta}{l^2} \frac{EI}{GA_c} \right]^{-1}. \tag{4.53}$$

Applying the principle of virtual work again, the rotated angle at the free-end due to the external force F_{ey} for the beam with and without wing structures are

$$\phi_{\text{wing}}^{F_{ey}} = \frac{F_{ey} l^2}{2EI} + \frac{F_{ey}(a_2 + 2a_3)b}{GJ}, \tag{4.54}$$

and

$$\phi_{\text{straight}}^{F_{ey}} = \frac{F_{ey} l^2}{2EI}, \tag{4.55}$$

respectively. Ratio $\zeta^{F_{ey}}$ between them is

$$\zeta^{F_{ey}} = \frac{\phi_{\text{wing}}^{F_{ey}}}{\phi_{\text{straight}}^{F_{ey}}} = 1 + 2\eta \frac{a_2 + 2a_3}{l} \frac{EI}{GJ}. \tag{4.56}$$

For a monomorph beam and assuming that $a_3 \gg a_1, a_2$ and $l \gg w, h$, Eq. 4.53 and Eq. 4.56 are simplified as

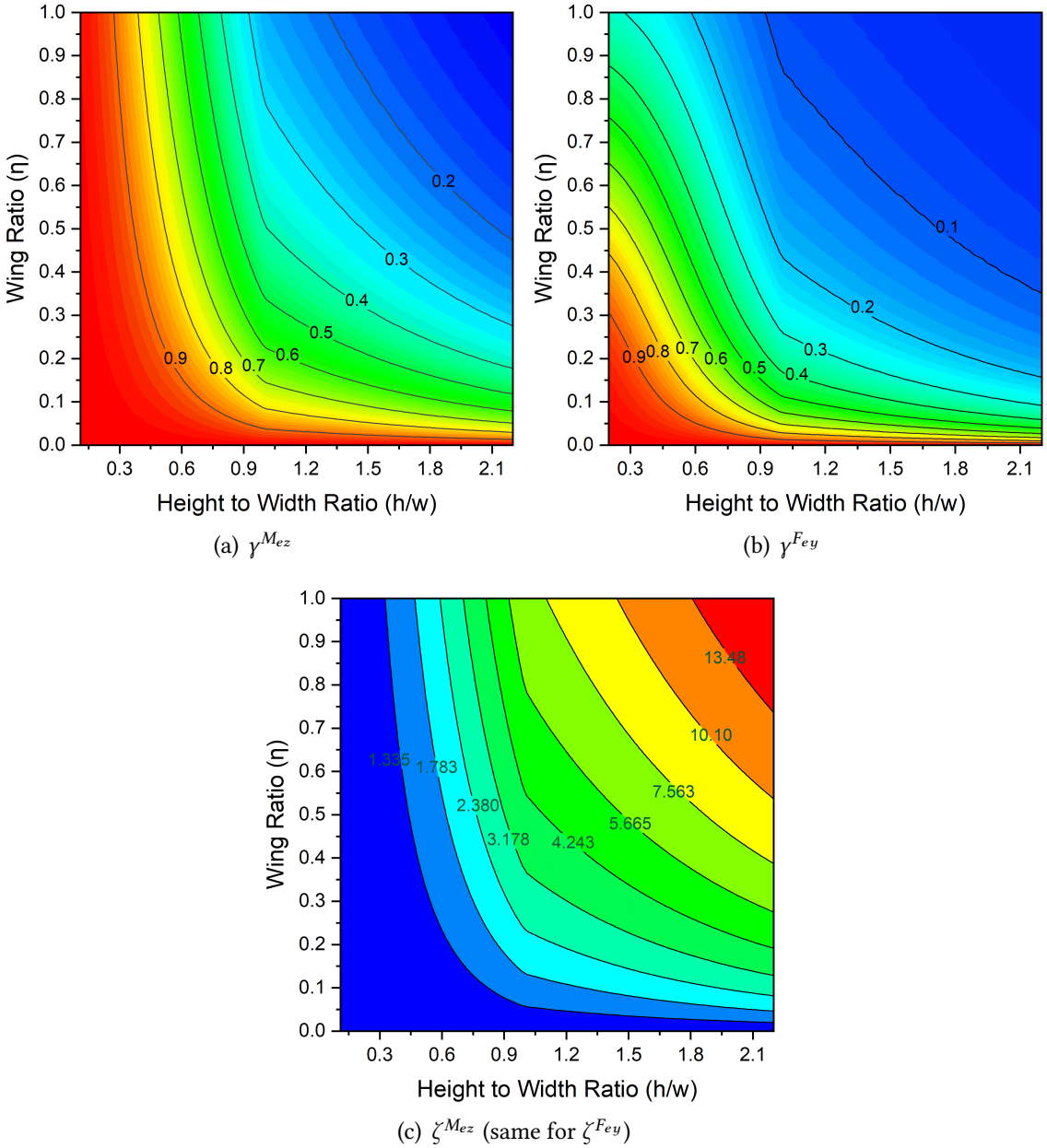
$$\gamma^{F_{ey}} \approx \left[1 + 2\eta^3 + \eta \frac{1+\mu}{\beta} \left(\frac{h}{w} \right)^2 \right]^{-1}, \tag{4.57}$$

and

$$\zeta^{F_{ey}} \approx 1 + \eta \frac{2(1+\mu)}{3\beta} \left(\frac{h}{w} \right)^2, \tag{4.58}$$

respectively.

Influences of η and ratio h/w on structural stiffnesses Based on Eqs. 4.49, 4.50, 4.57 and 4.58, it is interesting to learn that the structural stiffness is dependent on the external loads, although there is no explicit term of loadings in these equations. Besides, Eqs. 4.50 and 4.58 share the same expression. Fig. 4.9 demonstrates the influences of wing ratio η and height to width ratio h/w of the beam's cross-section on the structural stiffnesses under external loads M_{ez} and F_{ey}, respectively. The Poisson's ratio μ is set to 0.25.



(a) $\gamma^{M_{ez}}$ (b) $\gamma^{F_{ey}}$

(c) $\zeta^{M_{ez}}$ (same for $\zeta^{F_{ey}}$)

Figure 4.9.: Influences of wing ratio η and height to width ratio h/w of cross-section on structural stiffness ratio (a) $\gamma^{M_{ez}}$ (b) $\gamma^{F_{ey}}$ and (c) $\zeta^{M_{ez}}$ (same for $\zeta^{F_{ey}}$). Calculation based on half of monomorph double-beam structure in Fig. 4.7(b).

It is worth noting that here the range of h/w is chosen by taking the actual dimensions and limitations from fabrication into account. SOI chips with 220 nm device layer are used for device fabrication in studies of this dissertation, in other words, the beam height is fixed to be 220 nm. Due to limitations from lithography techniques, beam width is usually designed to be no less than 100 nm, which gives the upper limitation of height to width ratio $h/w < 2.2$. Meanwhile, to avoid severe underetching during releasing of Si beams, the width is usually chosen to be below 3 μm, which gives the lower limitation of $h/w > 0.07$.

Fig. 4.9(a) shows that η and h/w have almost the same influence on $\gamma^{M_{ez}}$, which decreases proportionally to increasing η and h/w. In the region of $h/w < 0.5$, the wing ratio has negligible effect on beam stiffness; however, in the region of $h/w > 1.2$, the stiffness is dramatically weakened by the increasing wing ratio. For range in between, the relation is almost linear. It shows a different trend if the beam undertakes shear force F_{ey}, as depicted in Fig. 4.9(b). The structural stiffness is more sensitive to η in the whole range studied of h/w, especially when h/w is larger than 0.9. Decreasing the stiffness is accompanied by a larger generalized displacement, which is demonstrated in Fig. 4.9(c). The trend in this figure is similar to the influence of η and h/w in Fig. 4.9(a).

As previously mentioned, Fig. 4.9 is plotted based on the monomorph beam that the bending stiffness EI and torsional rigidity GJ can be calculated conveniently. However, for a multilayer beam that Eqs. 4.45, 4.48, 4.53 and 4.56 need to be applied, the effective bending stiffness can still be acquired from Eq. 4.13, while deriving an accurate effective shear modulus is challenging [148]. Despite this, one can still use the concise law of mixture to calculate the effective shear modulus [148] and have the initial values of the parameters for structure design by exploiting this analytical model. A more accurate mechanical study can be achieved with finite element method (FEM) simulation. Nevertheless, these initial values help to speed up and facilitate the simulation process.

In short, one can select an appropriate wing ratio to have a desirable structural stiffness according to Fig. 4.9. On the one hand, a smaller stiffness arouses a larger deformation, which is usually beneficial for actuation; on the other hand, this also causes the structure to become fragile and "too sensitive" to external perturbation, which has been witnessed in fabrication.

4.4. Summary

In this chapter, an analytical model of multilayer beam structures is built based on Euler-Bernoulli beam theory to analyze the curvature and deflection after deformation due to inelastic strains in the layers. The equations obtained to calculate the deflection at the free-end of multilayer cantilever beam is applicable to scenarios of position-independent inelastic strains and large deflection. An analytical thermal model of Joule-heated multilayer beams is then built to obtain the temperature distribution, maximum temperature and power dissipation mechanism. It is found that heat conduction to the substrate plays

the major role in the heat dissipation mechanism and geometrical dimensions have negligible influence on the maximum temperature on the beam. This analytical model is also validated by the experiments of Joule-heating a cantilever beam at macroscale. The temperature profile of the beam captured by the infrared camera closely matches the profile predicted by the model after correction. A wing structure design has been introduced to overcome the large temperature gradient in Joule-heated beams in micro- and nano-scale. In the last section, the stiffness changes and induced displacement variations under two loading scenarios are investigated by using the energy method. Two important geometrical parameters, the wing ratio η and the height to width ratio of beam's cross-section h/w, are noticed and their influence on the structural stiffness is thoroughly studied. The corresponding criteria are proposed to have appropriate geometrical dimensions of wing structure designs, which will be investigated in Chapter 5.

5. Analyses of SMA/Si Bilayer Nanoactuators

In Chapter 4, the analytical mechanical and thermal models are built for the analyses of the general multilayer beam structures. SMA/Si bilayer cantilever beams are the most common and simplest type of SMA-based micro- and nanoactuators. Based on the built models, this chapter offers insights into the mechanical and thermal performances of the SMA/Si bilayer cantilever nanoactuators. Firstly, the deflection change of the bilayer beam with a homogeneous thermal cycle is studied to understand the influences of different effects on the mechanical behavior. Next, the phase transformation temperatures of a sputtered 600 nm-thick $Ti_{54}Ni_{31}Cu_{15}$ film on Si substrate are measured by exploiting the temperature-dependent resistance measurement, and the results are used in the thermal analysis of Joule-heating, including the temperature profiles and the power consumptions. The joint electro-thermo-mechanical analyses are performed afterwards to understand the detailed effect of the temperature gradient on the mechanical bending behavior. Several limitations and fundamental assumptions which can influence the analyzing results are discussed in the last section.

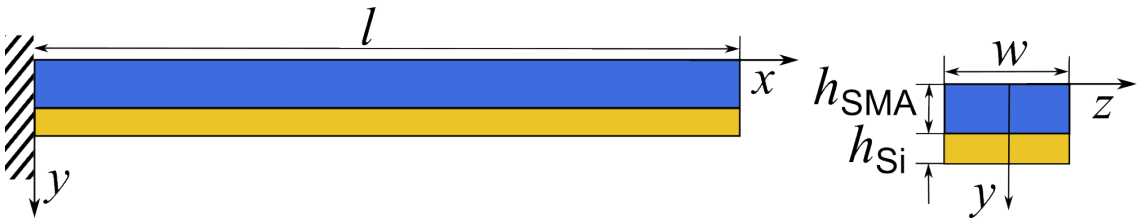
5.1. Mechanical Analysis

In Section 4.1, an analytical model describing the bending behavior of a multilayer beam due to inelastic strain has been proposed. Here, this model is used to analyze the deflection change of a SMA/Si bilayer cantilever beam across the phase transformation of SMA.

The structure to be analyzed is depicted in Fig. 5.1, which is composed by a SMA layer on the top and a Si layer at the bottom. The height of SMA and Si (h_{SMA} and h_{Si}) is 300 nm and 220 nm, respectively. The beam length l and width w are set to be 15 μm and 400 nm. Si is regarded as a linear elastic material with the detailed material properties listed in Table 5.1.

Table 5.1.: Material properties of Silicon.

Elastic Modulus	E_{Si}	170 GPa
Poisson's Ratio	μ_{Si}	0.28
Coefficient of Thermal Expansion	α_{Si}	$2.6 \times 10^{-6}\,°C^{-1}$



(a) Front View (b) Cross-section View

Figure 5.1.: Sketch of the analyzed SMA/Si bilayer cantilever beam.

Since the material properties of SMA are phase-dependent, it is necessary to introduce a temperature-dependent variable, the *volume fraction of martensite phase* $\xi(T)$, which helps to describe the phase transformation phenomenon of SMA. This variable is assumed to be in the form of a logistic function as

$$\xi(T) = \frac{1}{1 + \exp\left[k(T - T')\right]}, \tag{5.1}$$

where the terms k and T' are determined by the direction of phase transformation as below:

$$\text{Forward transformation:} \quad k = \frac{k_0}{M_s - M_f}, \quad T' = \frac{M_s + M_f}{2};$$

$$\text{Reverse transformation:} \quad k = \frac{k_0}{A_f - A_s}, \quad T' = \frac{A_s + A_f}{2}.$$

The constant k_0 is given in [149]. The parameters M_s and M_f refer to martensite start and finish temperatures, and A_s and A_f refer to austenite start and finish temperatures.

Therefore, the elastic modulus E_{SMA} and the coefficient of thermal expansion (CTE) α_{SMA} can be approximated by applying the law of mixture:

$$E_{\text{SMA}}(T) = E_m\xi(T) + E_a(1 - \xi(T)), \tag{5.2}$$

and

$$\alpha_{\text{SMA}}(T) = \alpha_m\xi(T) + \alpha_a(1 - \xi(T)), \tag{5.3}$$

where the subscripts m and a refer to the material properties at martensite and austenite phase, respectively.

The values of the material properties of the SMA layer used for analyzing are listed in Table 5.2. By applying Eqs. 4.1 – 4.5 and Eq. 4.9, the deflection change[1] at the free-end of the cantilever beam with temperature is plotted in Fig. 5.2(a). A clear hysteresis in the

[1] It is noteworthy that the deflection values in Fig.5.2 are scalar. The positive values do not contain the information of the bending direction of the beam.

curve is attributed to the phase transformation of SMA. Starting from room temperature (RT), the cantilever beam is already deflected owing to the tensile residual stress (σ^{re}_{SMA}) in the SMA layer. As the temperature in the beam increases until A_s, there is a small decrease of deflection due to the mismatch of CTE between SMA and Si. Thermal stress in the SMA layer slightly relieves the residual stress. When entering the phase transformation region, i.e., the temperature is above A_s, there is a larger and rapid increment in deflection, which can be explained through the following aspects. Firstly, transformation to Austenite phase hardens the material by having larger elastic modulus ($\Delta E-$effect); secondly, the martensite transformation strain gradually diminished during the reverse transformation [86]. After having fully transformed to austenite phase ($T > A_f$), the deflection curve becomes linear again as the effect of mismatch of CTE prevails again, which is similar to the situation below A_s. When the beam is cooled back to RT, the whole process reverses, except that the forward transformation takes place in the temperature range between M_s and M_f.

Table 5.2.: Material properties of Ti-Ni-Cu film adopted for the analytical model. Transformation temperatures and residual stress are measured by the project partner at Ruhr University Bochum. Data of elastic moduli are from [150], and CTE from [151].

M_s	62 °C	M_f	28 °C
A_s	37 °C	A_f	70 °C
E_m	41 GPa	E_a	71 GPa
α_m	15.4×10^{-6} °C^{-1}	α_a	15.4×10^{-6} °C^{-1}
σ^{re}_{SMA}	220 MPa	σ^{re}_{Si}	0 MPa

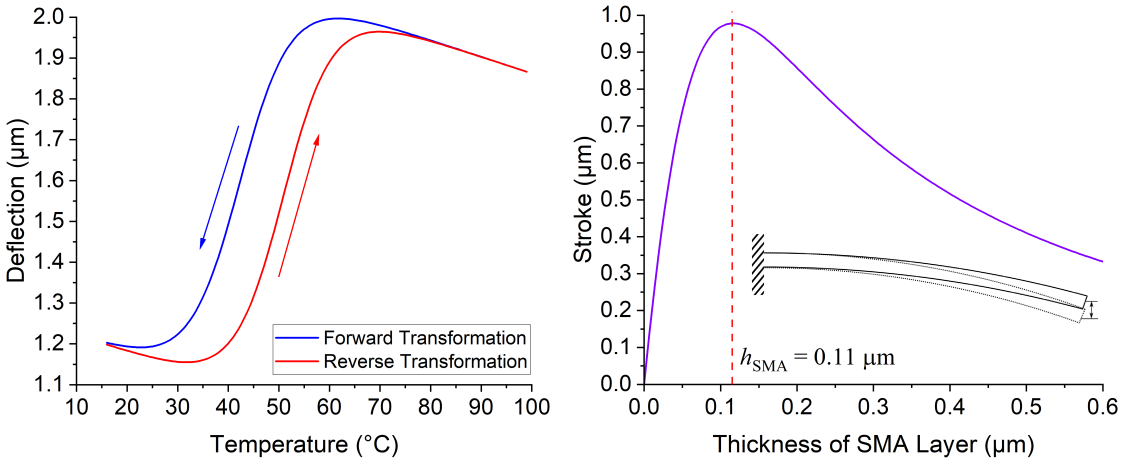


Figure 5.2.: Analytically calculated (a) deflection change of SMA/Si cantilever beam with temperature, and (b) SMA thickness-dependent stroke of the beam (thickness of Si layer fixed at 220 nm).

Because the material properties have been fixed in the fabrication process of SMA film and the thickness of Si is also determined by the SOI wafers used for fabrication, only the

tunable thickness of SMA is studied to learn its influence on the stroke of the cantilever beam in the following. The stroke is defined as the deflection difference between RT and 100 °C. Fig. 5.2(b) depicts its relationship with the thickness of SMA. The stroke increases fast and reaches a peak when the thickness of SMA increases from 0 to 0.11 μm. Afterwards, the stroke shows a declining trend as the thickness of SMA continues to grow. It can be inferred that the stroke will approach zero when $h_{\mathrm{SMA}} \gg h_{\mathrm{Si}}$ as the beam is structurally analogous to a monolayer beam. Therefore, it is advisable that the thickness of SMA is in the proximity of 0.11 μm when designing the bilayer cantilever beam in order to have large stroke.

5.2. Thermal Analysis

In Section 4.2, an analytical model has been built to demonstrate the large temperature gradient along micro- and nanocantilever and bridge beams when they are Joule-heated. Based on this model, a Joule-heated SMA/Si bilayer cantilever beam is investigated, and the temperature-dependent material properties are considered accordingly.

5.2.1. Temperature-dependent Material Properties

Material Properties of Si Resistivity ρ and thermal conductivity λ are two of the important material properties to consider when studying a Joule-heated structure. The resistivity of the p-type Si, ρ_{Si}, is $10\,\Omega\cdot\mathrm{cm}$ (at room temperature), which is provided by the manufacturer of the SOI wafers used for the fabrications in this dissertation. Since this value is far higher than the resistivity of SMA, and quantitatively describing the temperature-dependent resistivity of Si is complicated [152], ρ_{Si} is assumed to be a constant and irrelevant to temperature. However, the thermal conductivity of Si, $\lambda_{\mathrm{Si}}(T)$, is temperature-dependent and is expressed as

$$\lambda_{\mathrm{Si}}(T) = \frac{3.85 \times 10^4}{T} - 7.60 \times 10^{-8} T^{2.5}, \tag{5.4}$$

where the temperature T is in the unit of *Kelvin*. This semi-empirical formula is obtained by fitting the data from [153]. The inverse proportional term represents the approximation of Umklapp scattering process and the second term represents the approximation of phonon-electron scattering.

Material Properties of SMA To obtain the temperature-dependent resistivity of SMA, $\rho_{\mathrm{SMA}}(T)$, firstly a SMA film is sputtered on a Si substrate, which is then placed inside a cryostat to perform a temperature-dependent four-point measurement. The sputtering is conducted from an alloy target by the project partner at Ruhr University Bochum. The composition of the film is $\mathrm{Ti_{54}Ni_{31}Cu_{15}}$. Details of this measurement have been introduced in Section 2.1.4. Fig. 5.3 demonstrates the temperature-resistance curve of this 600 nm

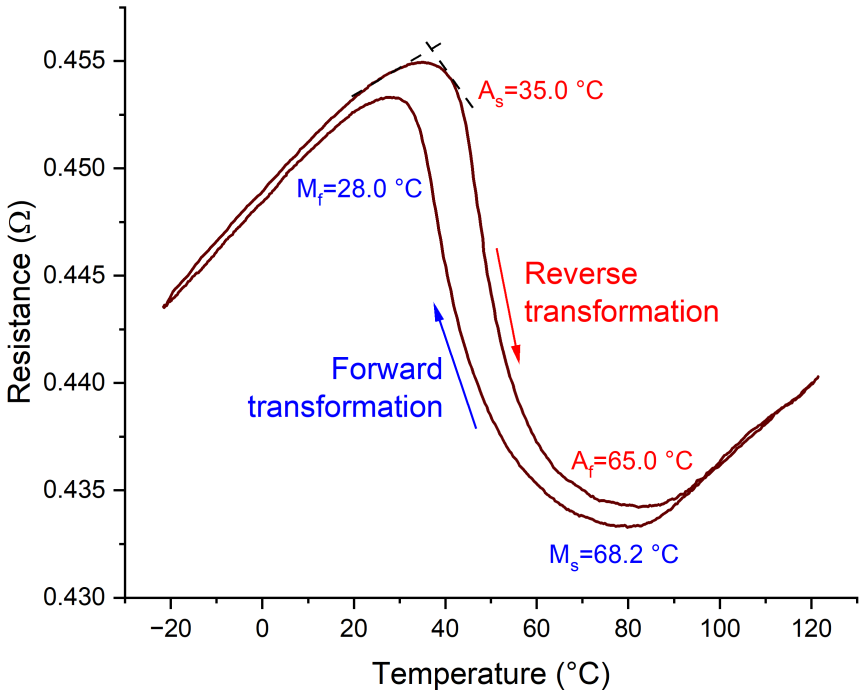


Figure 5.3.: Temperature-resistance curve of a sputtered 600 nm-thick $Ti_{54}Ni_{31}Cu_{15}$ film on Si substrate. Transformation temperatures marked in the figure are determined by the tangential method.

$Ti_{54}Ni_{31}Cu_{15}$ film, which shows the typical hysteresis during the phase transformation. The resistivity of the film is calculated based on the measured resistance data by Eq. 2.1.

The resistance shows a linear relation with temperature in the pure martensite and austenite regions ($T < M_f$ and $T > A_f$), while in the phase transformation region, the resistance change can still be described by the law of mixture. Therefore, the temperature-dependent resistivity of SMA is expressed as

$$\rho_{\mathrm{SMA}}(T) = \rho_m(T)\xi(T) + \rho_a(T)(1 - \xi(T)), \tag{5.5}$$

where the resistivity at martensite ($\rho_m(T)$) and at austenite ($\rho_a(T)$) are approximated by linear functions as

$$\rho_m(T) = \rho'_m T + \rho_{m,0}, \tag{5.6}$$

and

$$\rho_a(T) = \rho'_a T + \rho_{a,0}, \tag{5.7}$$

respectively. The constants ρ'_m and ρ'_a here are slopes of these functions, which can be obtained by fitting with the experimental data.

Based on Eq. 5.5, the temperature-dependent thermal conductivity of SMA is derived by applying the Wiedemann–Franz law as

$$\lambda_{\mathrm{SMA}}(T) = \frac{LT}{\rho_{\mathrm{SMA}}(T)}, \tag{5.8}$$

where L is the Lorenz number and equals to $2.44 \times 10^{-8}\,\mathrm{W\,\Omega\,K^{-2}}$. It is worth noting that the temperature in Eq. 5.8 should also be in *Kelvin*.

These temperature-dependent material properties will be adopted for the thermal analyses in the following. Due to the complexity of solving the differential Eq. 4.22, the analyses will be carried out numerically in Comsol. For comparison, constant values of these parameters at $T' = (A_s + A_f)/2$, e.g., $\rho_{\mathrm{SMA}}(T = T')$, are also used as solutions of the analytical model.

5.2.2. Temperature Profile and Power Consumption

The structure investigated is also the SMA/Si bilayer cantilever beam depicted in Fig 5.1 and the dimensions remain unchanged. A current is applied through the beam, and the temperature boundary conditions are given by Eq. 4.26. The surface convection coefficient is set to be $400\,\mathrm{W/(m^2\,K)}$. Fig. 5.4 shows the simulated and analytically solved profiles of the excess temperature[2] along the length direction of the beam. These two profiles show a close match with the maximum temperature difference about 1.5 K at the free-end of the beam.

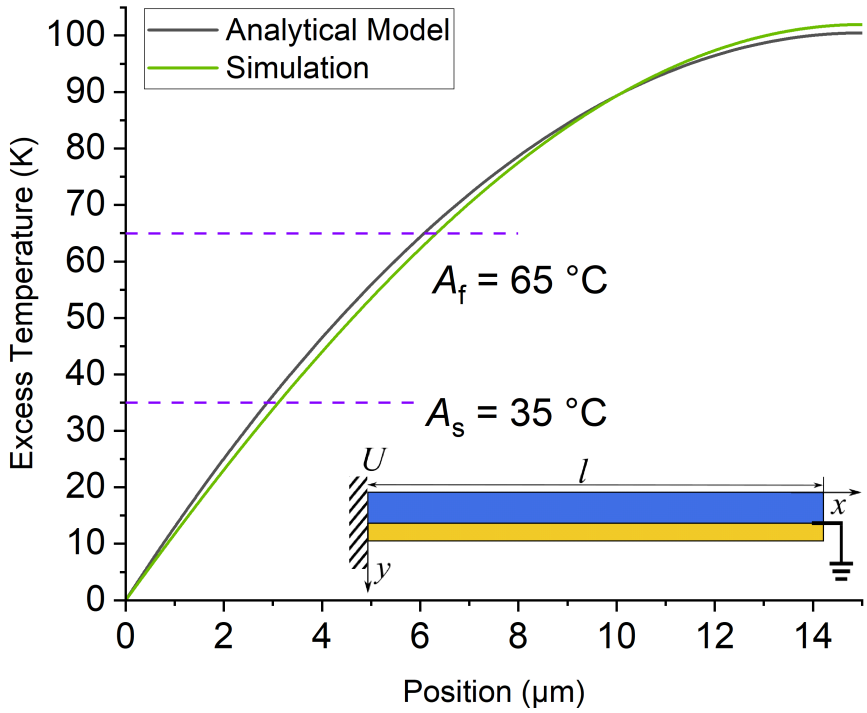


Figure 5.4.: Excess temperature profiles determined by an analytical model and by FEM simulation of a Joule-heated 15 μm long SMA/Si bilayer cantilever beam. The current applied is 1 mA.

Fig. 5.2(a) has revealed that the major deflection change of SMA/Si bilayer cantilever beam is contributed to the phase transformation, therefore, it is beneficial that the percentage

[2] The concept of excess temperature is defined by Eq. 4.27, which is the temperature difference between the actual temperature on the beam and the environmental temperature.

of beam length with $T > A_f$ is as high as possible. Fig. 5.5 illustrates the correlations between the increasing current and the rising percentage values that T exceeds A_s and A_f respectively. The current needs to be above a lower limit I_{min} that the temperature at the free-end of the beam is above the austenite finish temperature. As the current continues to rise, the percentage of beam length with $T > A_f$ increases rapidly at first, then gradually saturates and tends to 100%. Due to the boundary conditions that the fixed end is at room temperature, the beam cannot fully transform to austenite even though a large current is applied. Hence, it is necessary to compromise between the percentage value and the current applied. Here, the value 80% is chosen as most of the beam has transformed and further increasing of current has minor contribution to the growth of the percentage value.

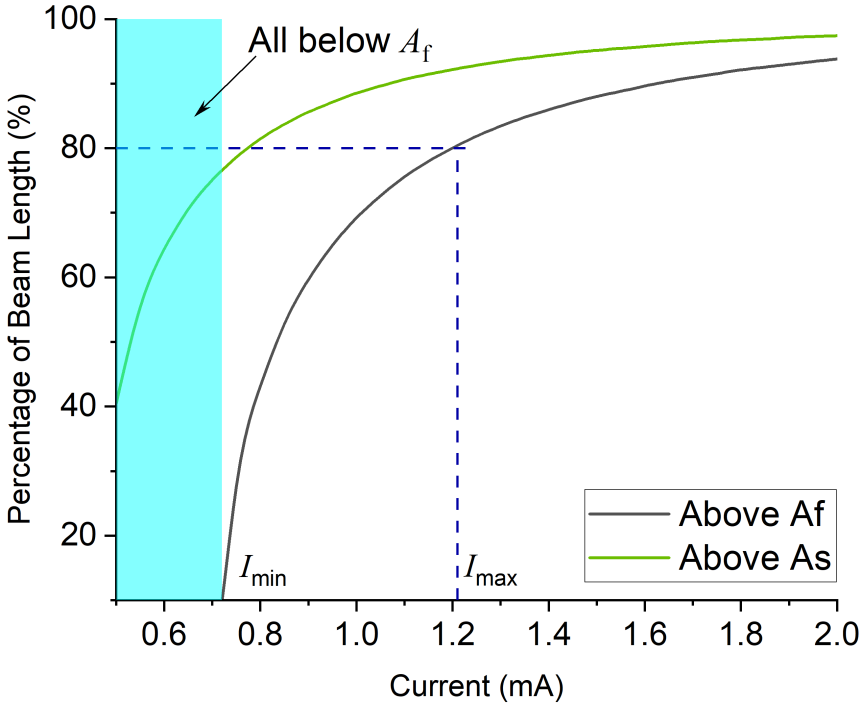


Figure 5.5.: The percentage of beam length starting ($T > A_s$) and finishing ($T > A_f$) phase transformation versus the heating current.

In Section 4.3, the wing structure design has been introduced for having a more homogenized temperature profile along the length direction of the beam. Thus, the section of the beam that is partially transformed can be folded into the wing structure. The mechanical softness of bending stiffness brought by the wing structure has been thoroughly studied in Section 4.3, and 20% of beam length folding into the wing leads to an acceptable decrease of bending stiffness by 40%–50%. This is another consideration of selecting the 80% criterion as the maximum current limit.

Power consumption and maximum temperature of the beam are the other two important aspects to investigate. Fig. 5.6 shows the relation between the current and these two parameters, which are obtained based on both the simulation and analytical models. There is minor difference between the two power curves, while the deviation between the two

curves of maximum excess temperatures, θ_{max}, enlarges as the current is increasing. The defined maximum current I_{max} corresponds to the excess temperature of about 160 K, which is also within a reasonable range that the intrinsic carriers of Si are not generated massively [154] thus avoiding the leakage of the current.

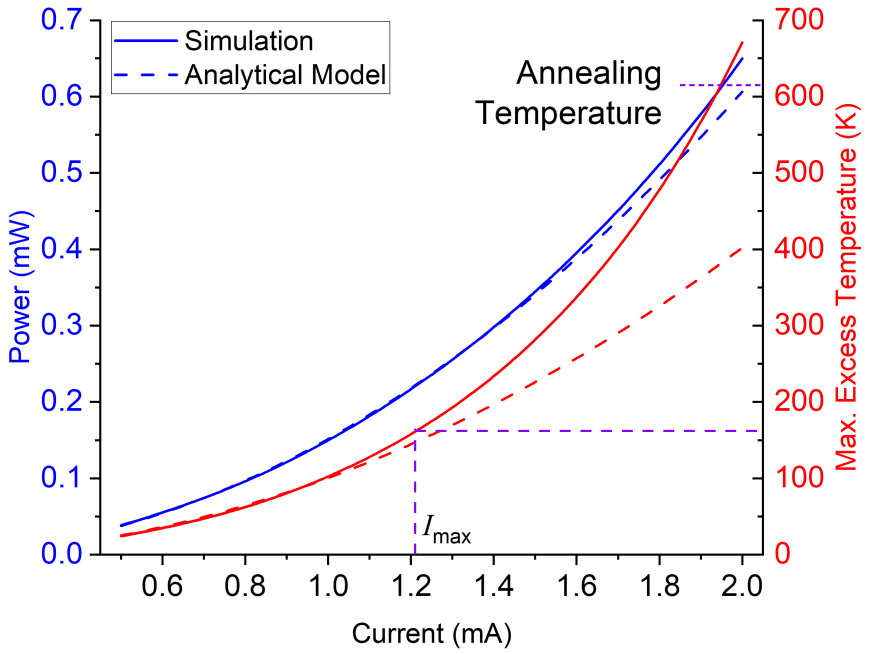


Figure 5.6.: Power consumption and maximum excess temperature versus the heating current.

5.3. Joint Thermo-mechanical Analysis

In Section 5.1, the mechanical analyses are based on that there is no temperature gradient along the beam, which has been proven to be inaccurate in the thermal analyses above. In this section, the electro-thermo-mechanical simulations are carried out in Comsol to investigate the thermo-mechanical performance of the SMA/Si bilayer cantilever beam. The design and boundary conditions remain the same as in the previous sections.

Fig. 5.7 demonstrates the deflection change of this 15 μm-bilayer cantilever beam with an increasing current applied, which ranges from 0.1 mA to 1.7 mA. When the current is below 0.45 mA, there is a minor deflection change as the maximum temperature on the beam is still below A_s and phase transformation has not started. A significant deflection increase can be witnessed when the current keeps rising and phase transformation takes place, which is similar to the trend in Fig. 5.2(a) based on the analytical mechanical model. However, after having surpassed the current that the maximum temperature is higher than A_f, i.e., the I_{min} defined in Section 5.2, the deflection is still growing due to the inhomogeneity of temperature along the beam, which is different from Fig. 5.2(a). The deflection reaches the peak at a current about 1.1 mA and starts to decrease afterwards, which can be explained by the competitive mechanism between the phase transformation and mismatch of CTE.

By referring to Fig. 5.5, it can be observed that about 75% length of the beam has finished the phase transformation at the current of 1.1 mA and further increase of current only contributes a small proportion of the beam to finish phase transformation. However, as discussed in Section 5.1, the mismatch of CTE dominates the deflection change when the temperature is above A_f and this effect leads to the decrease of deflection due to the relief of the residual stress. Apparently, the effect of mismatch of CTE prevails after the current exceeds the threshold value of 1.1 mA. In addition, this threshold is smaller than the $I_{\max}$ defined in Section 5.2.

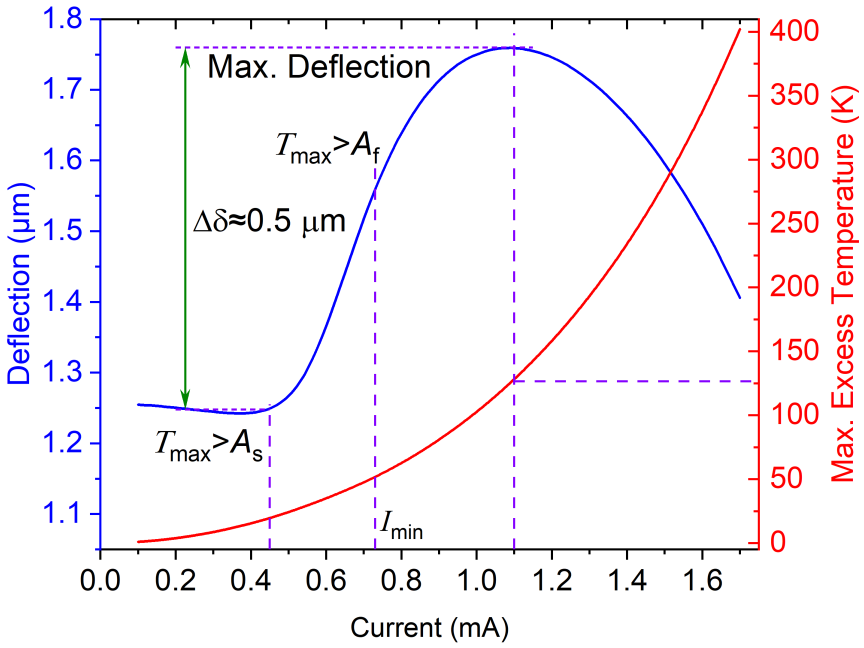


Figure 5.7.: Deflection variations of the Joule-heated SMA/Si bilayer cantilever beam with the current.

It is also introduced in Section 5.2 that 20% of the beam length can be folded into the wing structure to homogenize the temperature profile. The wing structure is further studied in Fig. 5.8, which demonstrates its influence on the deflections of the Joule-heated SMA/Si bilayer cantilever beams. The beam length is kept at 15 μm and the wing ratio η denotes the length ratio between the wing length and the beam length. Evidently, the increasing wing ratio elevates the maximum deflection value that the bilayer beam can reach, and the corresponding current required decreases. Taking the maximum deflection, and the corresponding maximum temperature and consumed power of the beam without wing structure as reference values, Fig. 5.9 illustrates the effects of the wing ratio on these three parameters. Although the increase of the deflection is not tremendous, both the corresponding maximum temperature and power witness considerable drops, which gradually level off when η is above 20%. This further confirms that 20%–25% is a suitable range that the wing ratio can be chosen from.

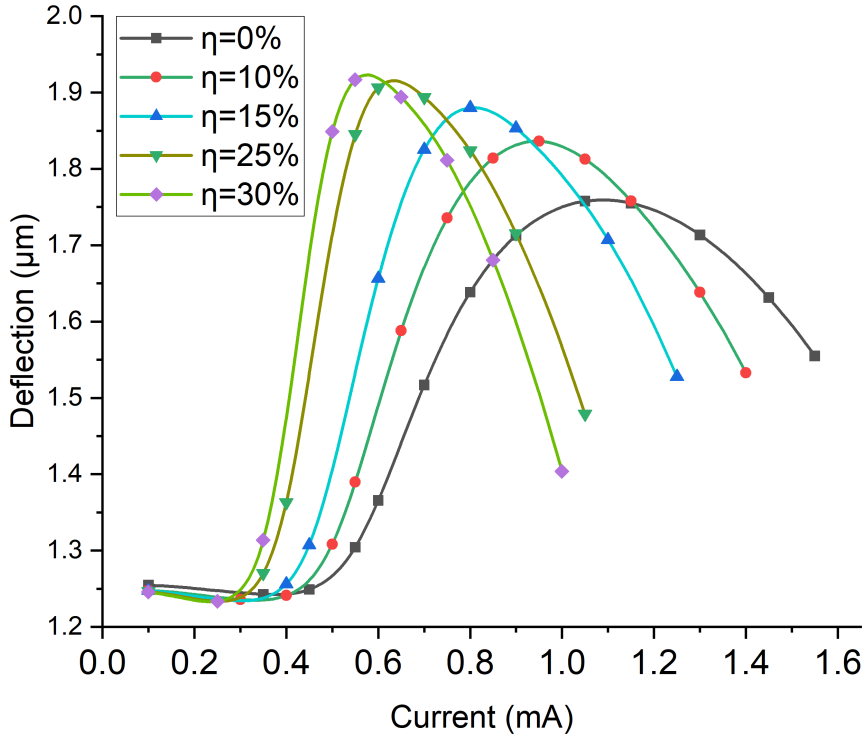


Figure 5.8.: Current-deflection curves of the Joule-heated SMA/Si bilayer cantilever beams with different wing ratio η. The beam length is kept at 15 μm.

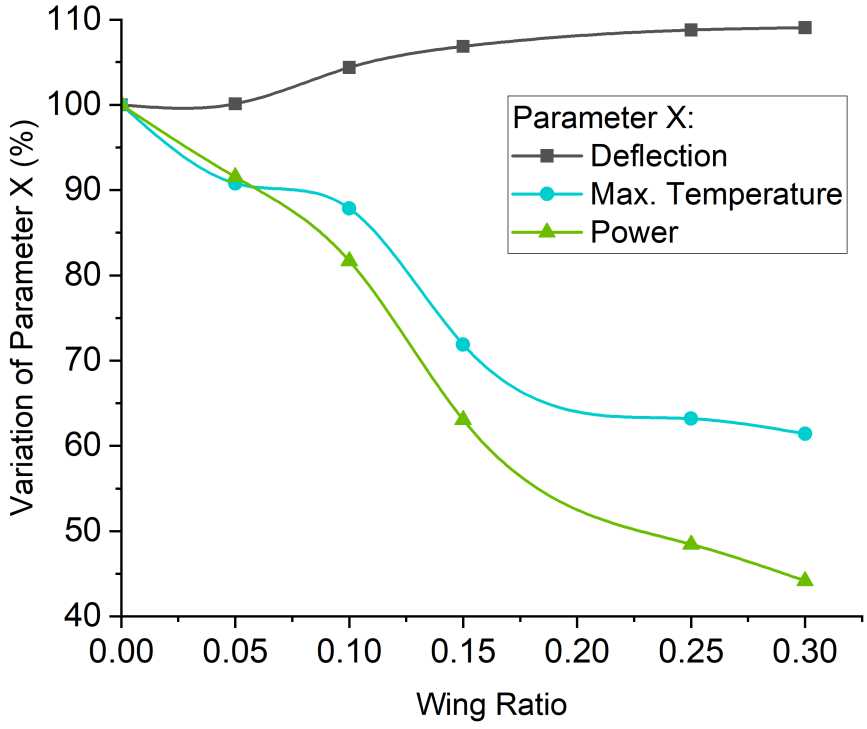


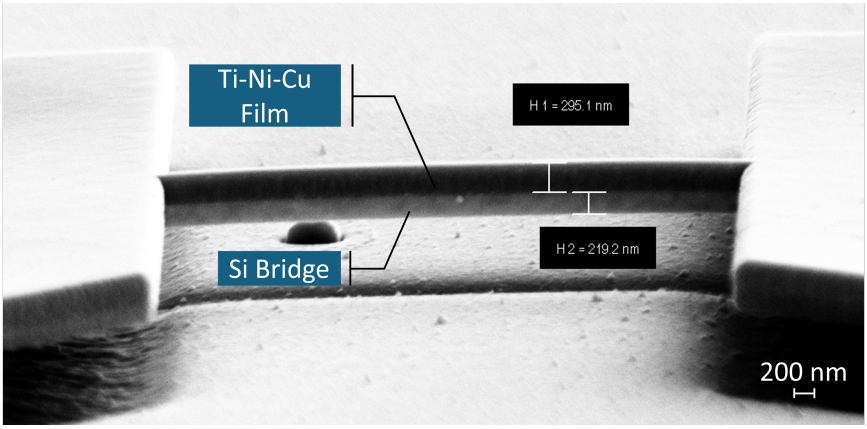
Figure 5.9.: Variations of deflection, maximum temperature and power with the increasing wing ratio. The values of the beam without wing are taken as references.

5.4. Discussion

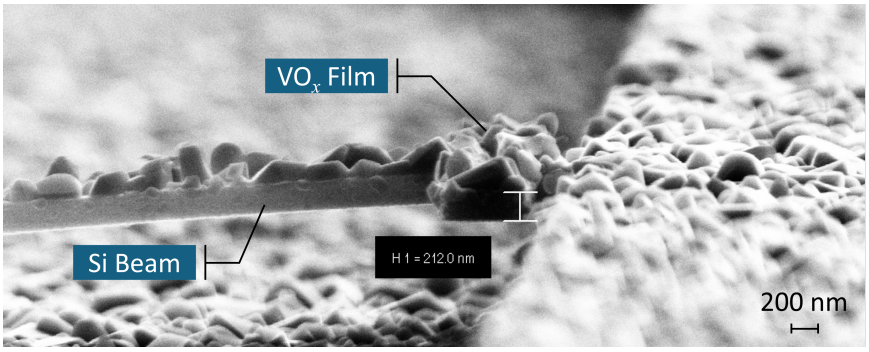
The above analytical and numeric models are developed to give straightforward insights into the mechanical, thermal and the joint performance of the Joule-heated SMA/Si bilayer cantilever beam nanoactuators. However, there are several factors that can cause deviations or even failures of these models, which will be discussed in the following.

5.4.1. Size-Effects

In Chapter 4 and this chapter, all the analytical and the numeric models are built within the frameworks of continuous medium, i.e., the medium can be continually divided into infinitesimal elements without separation. Usually, this assumption holds valid when the grain sizes of the functional film and the substrate are several orders smaller than the feature size of the structures. Fig. 5.10 presents two examples of different grain sizes. The sputtered Ti-Ni-Cu film in Fig. 5.10(a) is dense and smooth, whereas the sputtered VO_x film in Fig. 5.10(b) is discontinuous because of the coarse grain size. Apparently, the assumption of continuous medium is not applicable for the VO_x film in the picture.



(a) SMA film



(b) VO_x film

Figure 5.10.: Different films sputtered on free-standing Si beams: (a) dense Ti-Ni-Cu film with fine grain size, (b) discontinuous VO_x film with coarse grain size.

Another concern of the size-effect is the size-dependent material properties. Even though the films are dense, the material properties can vary with the thicknesses, which are the more common situations. For example, Curtis et al. reported that the phase transformation properties of Ti-Ni-Hf films are dependent on the film thicknesses and the types of substrate [37]. Chu et al. investigated the grain size effect on the elastic moduli of nanocrystalline Ni-Ti experimentally [155]. The influence of grain size effect on the hardness of nanocrystalline Ni-Ti was also studied [156]. König, et al. revealed that even the in-plane dimensions of structured Ti-Ni-Cu films influenced the phase transformation temperatures [157]. In these cases, the models based on the assumption of continuous medium remain valid while necessary correction terms are needed. Different theories have evolved to describe the size-dependent mechanical behavior at micro- and nanoscale as supplements for the classic continuum mechanics. For instance, the couple stress theory [158–164] introduces a material length scale parameter to describe this dependency, which was adopted to build the size-dependent constitutive model of SMAs by Choi et al. [165]. These theories describing the size-dependency can be further integrated into the current models.

5.4.2. Inhomogeneity of the Residual Stress

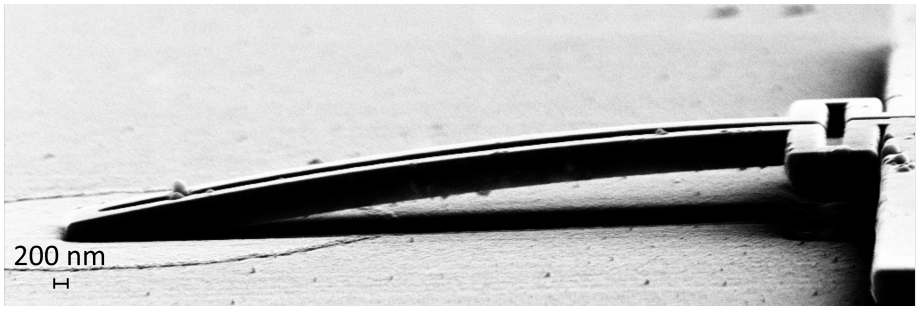
The residual stresses in the sputtered $Ti_{54}Ni_{31}Cu_{15}$ films are measured by monitoring the curvature changes of free-standing cantilever beams after being sputtered and calculating based on the Stoney equation (Eq. 4.21)[3]. The derivation of the Stoney equation presented in Section 4.1.3 indicates that the residual stresses in both the film and the substrate are assumed to be constant and irrespective of the position. However, due to the complexity of controlling the sputtering process, the homogeneity cannot always be guaranteed, as shown in Fig. 5.11. Both SMA/Si bilayer structures in Fig. 5.11 are on the same wafer but have different bending directions, indicating the variations of the residual stress across the wafer. This inhomogeneity may also cause the variations of the material properties, which leads to inaccurate results of the analytical and numeric models. Stress gradient can also exist in the substrate Si layer, which will be further discussed in Chapter 6.

5.5. Summary

In this chapter, the electro-thermo-mechanical performance of the Joule-heated SMA/Si bilayer cantilever nanoactuator is thoroughly investigated both analytically and numerically. The analytical mechanical model reveals the deflection evolution with hysteresis in a thermal cycle due to the competitions among several different effects, i.e., the CTE mismatch effect, the ΔE–effect and the shape memory effect. The phase transformation of a sputtered 600 nm-thick $Ti_{54}Ni_{31}Cu_{15}$ film on the Si substrate is characterized by the temperature-dependent resistance measurement, and the experimental results are used

[3] The measurements are performed by the project partners at Ruhr University Bochum. Further details of the technology can be found in [166].

(a) SMA/Si bilayer structure bending up.



(b) SMA/Si bilayer structure bending down.

Figure 5.11.: SMA/Si bilayer structures on the same chip show different bending directions: (a) bending-up, (b) bending-down.

for the electro-thermal analyses. Two delimitation values of the current, $I_{\min}$ and $I_{\max}$, are defined based on the temperature profiles of the bilayer beams, which guarantees that most of the SMA material can transform completely to the austenite state and the maximum temperature is within an acceptable range. Afterwards, a joint electro-thermo-mechanical study is performed on the bilayer beam structure. It is found that there is a peak of the deflection and the corresponding current is well below $I_{\max}$. Improvements of the deflection, the temperature gradient and the power consumption brought by the wing-structure design are also studied, and the suitable range of wing ratio of 20%–25% is further confirmed. There are situations, such as the size-effect and the inhomogeneity of the residual stress, that restrict the applications of the proposed models. Corrections of these situations can be the research focus in the future.

6. SMA Nanoactuators with Out-of-Plane Bistability

Bistable structures are well-known for their low energy consumption compared to other active structures as input energy is only needed during state switching. A mechanical bistable structure usually possesses two potential energy minima, which corresponds to two stable states. Such functionality is usually achieved by introducing a buckling mechanism due to its multi-solution features. Besides, there are external loads to switch between the two stable states. In this chapter, a bistable design is proposed by introducing the buckling of cantilever beams, which is induced by the internal residual stress in the SMA layer. Benefiting from the shape memory effect, the proposed design of bistable nanoactuators should actively switch between two buclked stable states. Important design criteria and parameters are investigated to ensure the bistability. The corresponding fabrication process and results are presented and discussed afterwards.

6.1. Classic Stability and Buckling Theory

Before introducing the bistable design, it is necessary to briefly review the classic stability and buckling theory. A slender beam can lose stability when it undertakes a compressive force above a certain critical value F_{cr}, which leads to dramatic increase in deflection value [139]. Such a phenomenon is named *buckling*.

As an example, a compressive force F is applied at the free-end of a slender cantilever beam with length l, which bends because of this force, as depicted in Fig. 6.1. Based on the Euler-Bernoulli beam theory and the assumption of small deflection, at position x the deflection is $y(x)$, and corresponding bending moment is

$$M(x) = F(y(l) - y(x)). \tag{6.1}$$

The relation between bending moment and curvature of a deflected beam is

$$\varphi(x) = \frac{M(x)}{EI}, \tag{6.2}$$

where EI is the bending stiffness of the beam [139]. Moreover, the curvature of plane curves is expressed as

$$\varphi(x) = \frac{y''(x)}{[1 + (y'(x))^2]^{3/2}}. \tag{6.3}$$

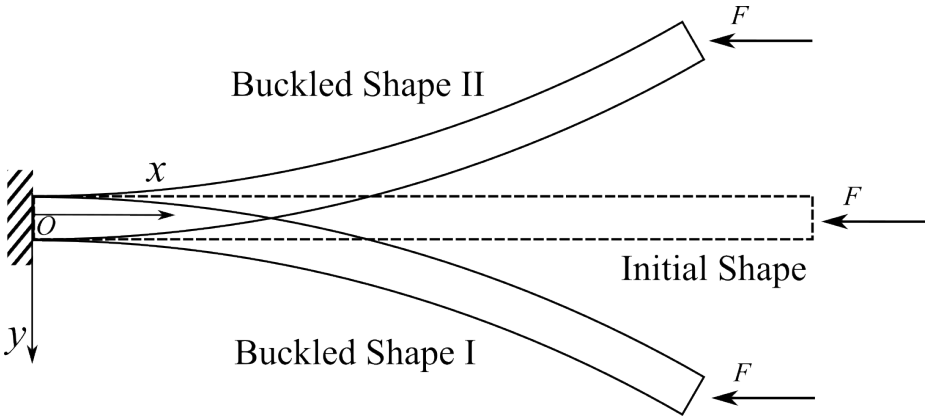


Figure 6.1.: A buckled cantilever beam under external compressive force F. The dashed line represents the initial shape and the solid line represents the two symmetric buckled shapes.

When the small deflection assumption holds, i.e., $y'(x)$ is small hence the term $(y'(x))^2$ is infinitesimally close to zero, combining Eqs. 6.1 – 6.3 leads to the bending equation of buckled beam

$$y''(x) = \frac{F(y(l) - y(x))}{EI}, \tag{6.4}$$

corresponding boundary conditions of the differential equation above are

$$y(0) = 0; \qquad y'(0) = 0. \tag{6.5}$$

Introducing variable k as

$$k^2 = \frac{F}{EI}, \tag{6.6}$$

solution of the differential equation, Eq. 6.4, based on Eq. 6.5 and 6.6 can be obtained as

$$y(x) = y(l)(1 - \cos kx), \tag{6.7}$$

with condition that k should satisfy

$$kl = \frac{n\pi}{2} \quad (n = 1, 3, 5, \cdots). \tag{6.8}$$

The deflection value at free-end of cantilever beam, $y(l)$, cannot be determined by equations above due to the assumption of small deflection and the simplification of curvature in

Eq. 6.4. Nevertheless, the critical force of buckling can be obtained from Eq. 6.6 and 6.8 as

$$F_{\mathrm{cr}} = \frac{\pi^2 EI}{(2l)^2}, \tag{6.9}$$

where n is taken to be 1, corresponding to the smallest value that F in Eq. 6.6 can take. It should be noted that, when n takes larger values, the values of F obtained correspond to the occurrence of higher modes of buckling. An intuitive understanding of high mode buckling based on Eq. 6.7 is that there is at least one deflection point in the deformed shape of buckled beam. More precise solutions of deformed shape rely on the accurate expression of curvature without adopting the small deflection assumption, which is based on the functions of elliptic integral [138] and is beyond the discussion here.

Furthermore, due to symmetry, there is another buckled shape drawn in Fig. 6.1. The corresponding equations of bending moment and deformed shape remain the same as Eq. 6.1 and Eq. 6.7 in the same coordinate system. Based on Eq. 4.38 which calculates the strain energy of the bent beam, apparently these two deformed shapes lead to the same strain energy. Therefore, it proves that a buckled beam possesses two buckled shapes with identical energy level. Initial disturbance during loading and structural imperfection determine the real shape that the beam actually buckles to. This forms the theoretical foundation of creating a mechanical bistable structure by buckling a beam structure.

In a more general case that is not limited to cantilever beam, the critical force can be written as:

$$F_{\mathrm{cr}} = \frac{\pi^2 EI}{(\mu l)^2} \tag{6.10}$$

where μ is named as *effective length factor* and is dependent on the boundary conditions of beam. For example, in case of a beam fixed at both ends, μ equals to 0.5; for a beam fixed at one end and pinned at the other end, μ approximately equals to 0.7.

The corresponding critical stress on the cross-section that leads to buckling of the beam structure is

$$\sigma_{\mathrm{cr}} = \frac{\pi^2 EI}{(\mu l)^2 A_c}, \tag{6.11}$$

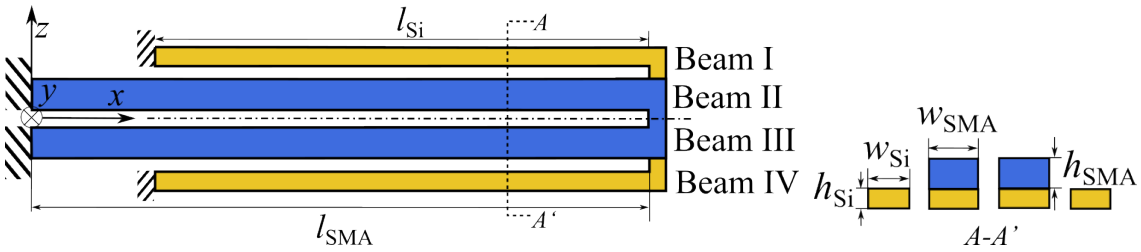
where A_c is the area of the cross-section.

6.2. Structural Design of the Nanoactuators

6.2.1. Bistable Mechanism

As introduced above, the key point to have a bistable beam structure is to buckle it with a compressive force that is larger than the critical force. Considering that the residual stress in annealed Ti-Ni-Cu films is a positive tensile stress [86], and assuming that there is no stress in the device layer of SOI wafers used for fabrication, a design of four cantilever beams connected in parallel is proposed as depicted in Fig. 6.2.

For a clear description of the bistable motion mechanism, the four cantilever beams are named as Beam I to IV, shown in Fig. 6.2. The common substrate of all four beams is Si and the Beam II and III are locally coated with a SMA layer. The overall structure is designed to be symmetric in order to overcome the in-plane rotation torsion due to the deformation of beams, namely, Beam I and IV, Beam II and III are identical, respectively. Four necessary geometrical parameters are introduced to define the beams: the widths of Beam I and II (w_{Si} and w_{SMA}), and the lengths of Beam I and II (l_{Si} and l_{SMA}), which are marked out in Fig. 6.2. Beam II and III form a current circuit for Joule-heating of the SMA layer on these two beams.



(a) Top view (b) Section view *A-A'*

Figure 6.2.: (a) Top view and (b) section view *A-A'* of the designed bistable beam structure. The two beams in the middle, Beam II and III, are locally coated with a SMA layer.

Due to the tensile residual stress in the SMA layer, the potential to "shrink" acts as a compressive force on Beam I and IV, which is able to buckle them when this force is larger than the critical force. Once these two beams are buckled, they possess two stable states, which are either buckled up or down.

Switching between the two stable states (State I and II) of buckled Beam I and IV is achieved by applying different heating cycles to the SMA layer and controlling the highest temperature of the cycle whether it is below A_s or above A_f. In detail, starting from room temperature in State I by applying a first heating cycle until a certain temperature $T < A_s$ is reached and cooling back to room temperature, Beam I and IV should be able to switch to the second stable buckling state (State II) due to the different CTE effect. In a second heating cycle, the highest temperature should be above the austenite finish temperature A_f so that the phase transformation can finish. Due to the shape memory effect, the SMA layer push the Beam I and IV back to their original stable state (State I). This accomplishes the reversible switching between the two stable states of the structure without introducing external loads.

Meanwhile, there are some important details to discuss in the following to ensure the bistability of the four-beam nanoactuator proposed above. Due to symmetry, half of the structure, i.e., Beam I and II, is taken for the subsequent analyses, as shown in Fig. 6.3.

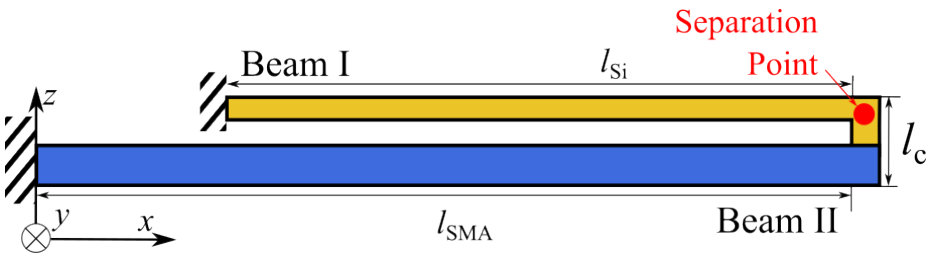


Figure 6.3.: Only half of the four-beam structure is considered for analyses due to symmetry. Beam I and II are separated at the marked separation point for analyses in the following.

6.2.2. The Compressive Force and Critical Force of Buckling

Originating from the inelastic strain of the SMA layer, the compressive force from Beam II can be calculated by Eq. 4.15 as an equivalent force denoted as F_{eq}. The critical force F_{cr} to buckle Beam I is defined by Eq. 6.10. Meanwhile, a buckling mode higher than one should be avoided, which can lead to falling into local energy minima of higher modes. Therefore, the criterion between these two forces is

$$F_{\text{cr}}^{\text{1st}} < F_{\text{eq}} < F_{\text{cr}}^{\text{2nd}}, \tag{6.12}$$

where the superscript of F_{cr} represents the buckling mode.

However, a key point to determine F_{cr} by Eq. 6.10 is the value of μ, which is dependent on the boundary conditions (or supporting conditions) of the buckled beam. Here, the equivalent procedure is taken that Beam I and II in Fig. 6.3 are separated and the buckled Beam I is regarded as a cantilever beam on an elastic support defined by Beam II, which provides three spring stiffnesses (k_x, k_y, α_r) in $x-$, $y-$direction and rotation respectively, as shown in Fig. 6.4. It is worth noting that though Beam II indeed also provides spring stiffnesses in other direction, for instance, rotation in $x-$direction, only the three stiffness variables considered above can influence deflection of Beam I in $y-$direction.

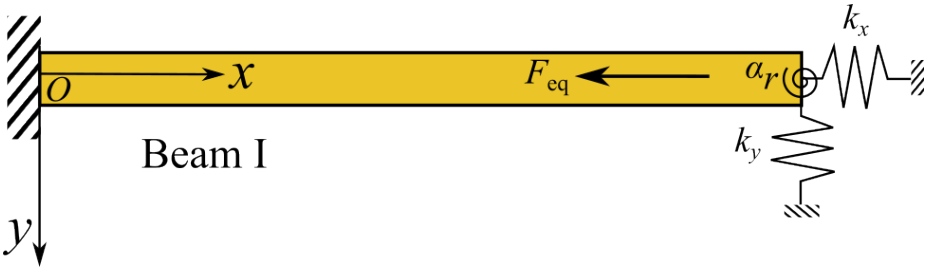


Figure 6.4.: Beam I simplified as a cantilever beam on an elastic support with spring stiffnesses k_x, k_y and α_r, which are provided by Beam II (*Side view* of Fig. 6.3).

As demonstrated in Section 4.3, structural stiffness can be obtained by applying Castigliano's second theorem, which is also followed here to calculate the stiffness of the elastic support from the remaining structures in Fig. 6.3. As depicted in Fig 6.5, different types of virtual external loads corresponding to the three spring stiffnesses in Fig. 6.4 are applied at the separation point. The boundary condition of this separation point is another

challenging point because it is also elastically connected to Beam I, in other words, it is in a situation between free-standing and fixed. In the following, the stiffnesses k_x, k_y, α_r are calculated in these two scenarios, respectively. Thereby, it is assumed that the actual values of these stiffnesses should fall into the range defined by the two extreme scenarios.

Figure 6.5.: Beam II with virtual external loads at the separation point to calculate its structural stiffness.

Free-standing boundary condition at separation point: The spring stiffnesses are calculated when the loads in Fig 6.5 are applied separately and their expressions are

$$k_x = \left[\frac{l_{\mathrm{SMA}}}{(EA)_{\mathrm{SMA}}} + \frac{l_c^3}{3(EI)_{\mathrm{Si,\,y}}} + \frac{k l_c}{(GA_c)_{\mathrm{Si}}} + \frac{l_{\mathrm{SMA}} l_c^2}{(EI)_{\mathrm{SMA,\,y}}} \right]^{-1}, \tag{6.13}$$

$$k_y = \left[\frac{l_{\mathrm{SMA}}^3}{3(EI)_{\mathrm{SMA,\,z}}} + \frac{k l_{\mathrm{SMA}}}{(GA_c)_{\mathrm{SMA}}} + \frac{l_c^3}{3(EI)_{\mathrm{Si,\,z}}} + \frac{k l_c}{(GA_c)_{\mathrm{Si}}} + \frac{l_{\mathrm{SMA}} l_c^2}{(GJ)_{\mathrm{SMA}}} \right]^{-1}, \tag{6.14}$$

and

$$\alpha_r = \left[\frac{l_{\mathrm{SMA}}}{(EI)_{\mathrm{SMA,\,z}}} + \frac{l_c}{(GJ)_{\mathrm{Si}}} \right]^{-1}, \tag{6.15}$$

respectively. l_c is the gap size between Beam I and II, as marked in Fig. 6.3. The subscripts "Si" and "SMA" refer to the type of the cross-sections of Beam I and Beam II, respectively. The subscripts "y" and "z" represent the reference axis for calculating the area moment of inertia I. The factor k is given in Eq. 4.39.

Fixed boundary condition at separation point: The loads in Fig 6.5 are applied simultaneously, therefore the generalized displacements induced by these loads can be obtained based on Eqs. 4.36 – 4.41, which are

$$\delta_x = \frac{\partial U(F_x, F_y, M_z)}{\partial F_x}, \quad \delta_y = \frac{\partial U(F_x, F_y, M_z)}{\partial F_y}, \quad \phi_z = \frac{\partial U(F_x, F_y, M_z)}{\partial M_z}. \tag{6.16}$$

The spring stiffnesses are obtained by letting two displacements in Eq. 6.16 to zero. For instance, k_x is calculated by

$$k_x = \frac{F_x}{\delta_x} \qquad \text{with conditions} \quad \delta_y = 0 \quad \text{and} \quad \phi_z = 0. \tag{6.17}$$

The spring stiffnesses therefore are:

k_x remains the same as Eq. 6.13.

$$k_y = \left[\left(1 - \frac{3M_z}{2F_y l_{\text{SMA}}} \right) \frac{l_{\text{SMA}}^3}{3(EI)_{\text{SMA, z}}} + \frac{k l_{\text{SMA}}}{(GA_c)_{\text{SMA}}} + \frac{l_c^3}{3(EI)_{\text{Si, z}}} + \frac{k l_c}{(GA_c)_{\text{Si}}} + \frac{l_{\text{SMA}} l_c^2}{(GJ)_{\text{SMA}}} \right]^{-1}, \tag{6.18}$$

where

$$\frac{M_z}{F_y l_{\text{SMA}}} = \left[2 + 2 \frac{l_c}{l_{\text{SMA}}} \frac{(EI)_{\text{SMA, z}}}{(GJ)_{\text{Si}}} \right]^{-1}.$$

Because the gap between Beam I and II, l_c, is usually much smaller than the beam length, i.e., $l_c \ll l_{\text{SMA}}$, the ratio term $M_z / F_y l_{\text{SMA}}$ approximately equals to $1/2$. And the rotational stiffness is

$$\alpha_r = \left[\left(1 - \frac{F_y l_{\text{SMA}}}{2M_z} \right) \frac{l_{\text{SMA}}}{(EI)_{\text{SMA, z}}} + \frac{l_c}{(GJ)_{\text{Si}}} \right]^{-1}, \tag{6.19}$$

where

$$\begin{aligned} \frac{F_y l_{\text{SMA}}}{M_z} = & \left[\frac{2}{3} + \frac{2k}{l_{\text{SMA}}^2} \frac{(EI)_{\text{SMA, z}}}{(GA)_{\text{SMA}}} + \frac{2}{3} \left(\frac{l_c}{l_{\text{SMA}}} \right)^3 \frac{(EI)_{\text{SMA, z}}}{(EI)_{\text{Si, z}}} + \frac{2k l_c}{l_{\text{SMA}}^3} \frac{(EI)_{\text{SMA, z}}}{(GA_c)_{\text{Si}}} \right. \\ & \left. + 2 \left(\frac{l_c}{l_{\text{SMA}}} \right)^2 \frac{(EI)_{\text{SMA, z}}}{(GJ)_{\text{SMA}}} \right]^{-1}. \end{aligned}$$

Similarly, taking the approximation that l_c / l_{SMA} is close to zero, the ratio term $F_y l_{\text{SMA}} / M_z$ is simplified as

$$\frac{F_y l_{\text{SMA}}}{M_z} \approx \left[\frac{2}{3} + \frac{2k}{l_{\text{SMA}}^2} \frac{(EI)_{\text{SMA, z}}}{(GA_c)_{\text{SMA}}} \right]^{-1}.$$

Eqs. 6.13 – 6.15 and Eqs. 6.18 – 6.19 define the lower and upper bounds that the values of the spring stiffnesses (k_x, k_y and α_r) can take. It can be noticed that the main difference between these two scenarios is the factor of the term relating to the length and cross-section property of Beam II, i.e., the beam with SMA layer.

To determine the critical force of buckling of the beam in Fig. 6.4, Michael and Kwok [49] have derived the value of effective length factor μ when k_y is infinite. Unfortunately, their method is not applicable when k_y is finite due to the concern of complexity of solving high-order differential equations. Therefore, the buckling behavior of the structure in Fig. 6.4 is simulated in Comsol with the calculated spring stiffnesses to investigate the critical force to ensure the validity of Eq. 6.12.

6.2.3. The Moment for State-switching

To have a first impression of the switching of stable states of Beam I in Fig. 6.4, the pure Si structure with the parameters in Table 6.1 is simulated in Comsol. Firstly, the first and second mode of the critical forces of buckling are calculated to be 13.56 µN and 30.52 µN, respectively. Afterwards, a displacement load in $y-$direction is applied at its right free-end and the reaction moment also at the free-end is extracted from the simulation results. The relation between the displacement and the reaction moment is plotted in Fig. 6.6, which demonstrates a typical bistable behavior. The curve crosses the zero moment line for three times, which represents the two stable states and the unstable state during state-switching. The maximum moment of the curve denotes the threshold of moment for successful state-switching.

Table 6.1.: Values of parameters adopted for the illustrative example of the bistable behavior of buckled Beam I.

w[1]/µm	h/µm	l_{Si}/µm	k_x/N · m^{-1}	k_y/N · m^{-1}	α_r/µN·µm	F/µN
0.7	0.22	15	1659.29	0.24	72.70	23.73

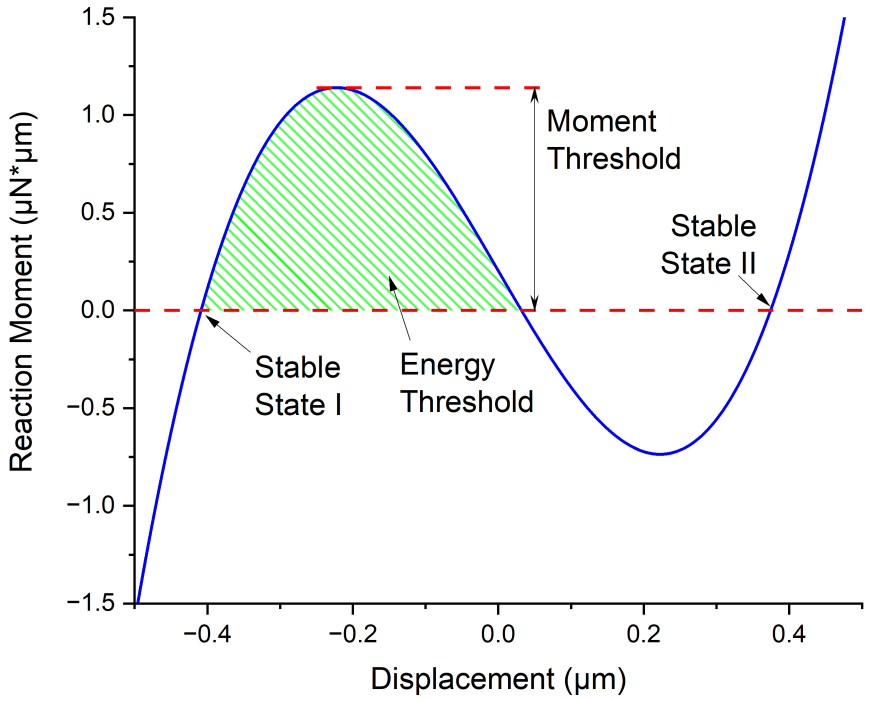


Figure 6.6.: Simulated reaction moment-displacement curve of buckled Beam I showing bistability. Parameters used for simulation are listed in Table 6.1. Shaded area represents the energy threshold that needs to be overcome for state-switching.

[1] w and h here represent the width and the height of the cross-section of beam, respectively.

However, from Eqs. 4.15 and 4.17, it can be learned that the residual stress in the SMA layer not only contributes to the force for buckling in Beam I, but also results in a "residual moment", which is unfavorable for the state switching. The active thermal moment generated in the SMA/Si bilayer of Beam II needs to be large enough to overcome this residual moment plus the moment threshold of state-switching, which gives another limit on the design of dimensions.

Overall, one can conclude that the geometrical dimensions of this four-beam structure interplay with each other in a complicated way:

- The critical force of buckling of Beam I is heavily dependent on the dimensions of itself and Beam II (through spring stiffnesses k_x, k_y and α_r).
- The actual compressive force for buckling is determined by the residual stress and dimensions of SMA layer on Beam II and the force should be larger than the critical force of buckling.
- Similar to the compressive force, the thermal moment is dependent on the residual stress and the dimensions of the SMA layer on Beam II as well, whereas the moment threshold is also influenced by the compressive force.

6.3. Parametric Study

In Section 6.2, the influences of geometrical dimensions on structural performance have been explained qualitatively. In this section, these influences are explored quantitatively through parametric study, which is beneficial for discovering the design space of geometrical dimensions of the proposed four-beam structure.

6.3.1. Spring Stiffnesses

Eqs. 6.13 – 6.15 and Eqs. 6.18 – 6.19 have revealed the relations among the spring stiffnesses, material properties and geometrical dimensions of Beam II. Meanwhile, there are parameters that have been constrained from the view of fabrication, including the elastic moduli and thicknesses of the Si and SMA layers. Besides, one can notice that the width of Beam II, w_{SMA}, has a simple linear relation with these three stiffness values, hence the influence of width is not studied here. The gap size, l_c, is also negligible as it is far smaller than the lengths of both Beam I and II. Therefore, the length of Beam II, l_{SMA}, is the focus of parametric investigation on the spring stiffnesses in the following.

The values of parameters used for analyzing are listed in Table 6.2. Based on Eqs. 6.13 – 6.15 and Eqs. 6.18 – 6.19, Fig. 6.7 demonstrates the variations of the three spring stiffnesses, k_x, k_y and α_r, with the length l_{SMA} ranging from 5 μm to 40 μm. Generally, they all have a inversely proportional relation with l_{SMA}. For k_y and α_r having the upper and lower bounds, the difference between the two bounds also shrinks rapidly as l_{SMA} increases. The

average values of the two bounds of k_y and α_r are taken for further analyzing hereinafter, and for this reason, larger l_{SMA} is preferred to reduce the error of using the average value.

Table 6.2.: Values of the parameters adopted for analyzing the spring stiffnesses induced by the rigidity of Beam II.

E_{Si}/GPa	h_{Si}/µm	E_{SMA}/GPa	h_{SMA}/µm	w_{SMA}/µm
170	0.22	41	0.3	1.0

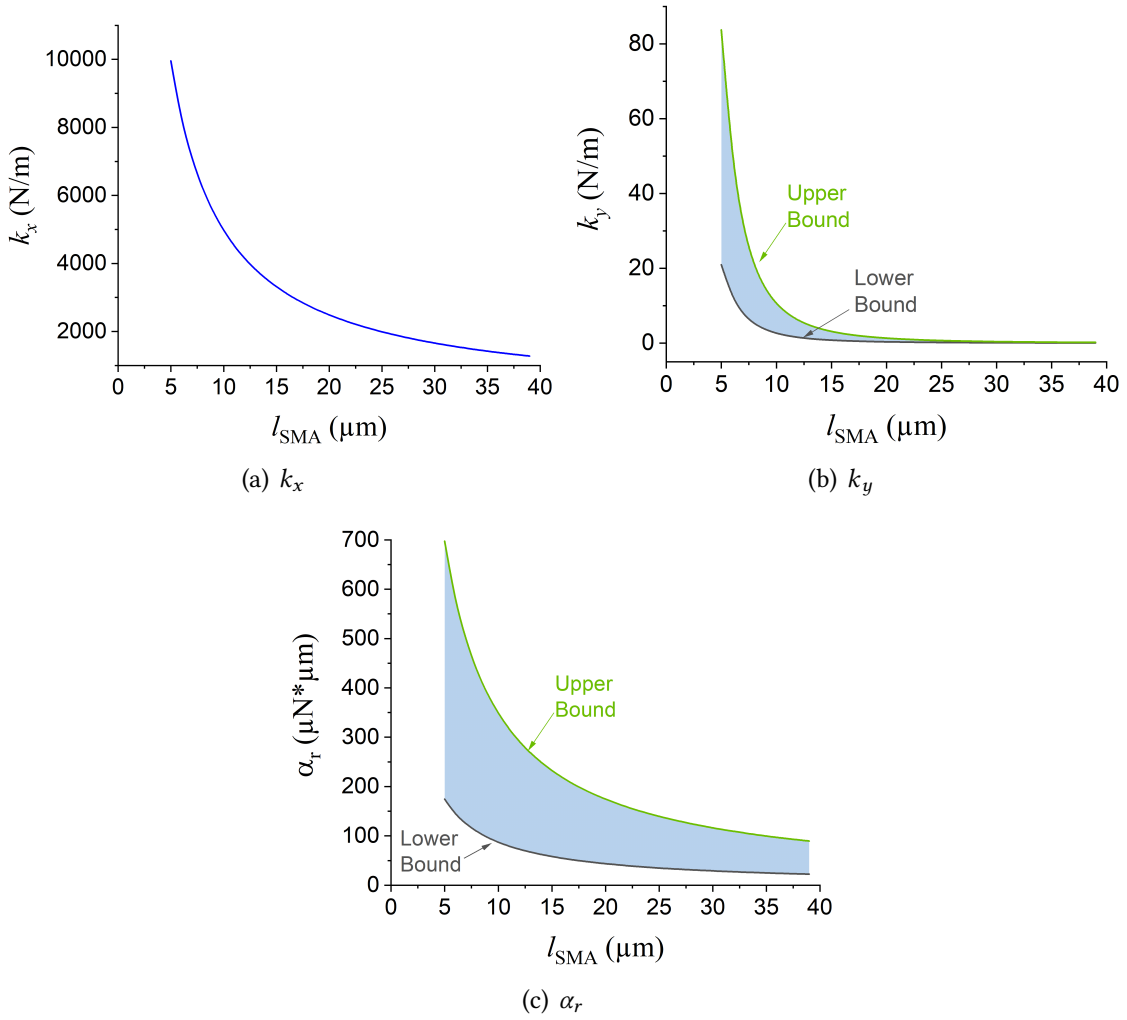


Figure 6.7.: Influence of the length of Beam II l_{SMA} on the spring stiffnesses (a) k_x, (b) k_y and (c) α_r. The shaded areas in (b) and (c) define the regions that the actual spring stiffnesses should fall into.

6.3.2. Critical Forces of Buckling

As introduced previously, the critical forces of buckling of Beam I can be numerically calculated in Comsol once the three spring stiffness values are known. In the following

the interactions between the critical forces and different parameters are investigated and visualized, which includes the length of Beam II l_{SMA} (Fig. 6.8), the length of Beam I l_{Si} (Fig. 6.9) and the width of Beam I w_{Si} (Fig. 6.10). Values of the parameters used for all the studies in this section are listed in Table 6.2 and Table 6.3, except that the parameter is chosen for investigation.

Table 6.3.: Values of the parameters adopted for analyzing the critical forces of buckling.

$l_{\mathrm{Si}}/\mu\mathrm{m}$	$w_{\mathrm{Si}}/\mu\mathrm{m}$	$l_{\mathrm{SMA}}/\mu\mathrm{m}$
15	0.7	30

Length of Beam II l_{SMA}: It has already been illustrated in Fig. 6.7 that increasing l_{SMA} leads to "softer" spring stiffnesses. In other words, constraints on the bending of the beam from the springs are smaller. Fig. 6.8 further confirms this softening by revealing that both the first and second mode of critical forces of buckling decrease with increasing l_{SMA}. Worthy of particular attention is that the curves of these two forces approach to each other at first, then separate again before crossing. Afterwards, the difference gradually saturates though l_{SMA} is still growing. The length corresponding to the smallest difference of first and second mode of critical forces is 18.5 μm. By comparing the buckling shapes at different l_{SMA}, as drawn in Table 6.4, it can be noticed that there is a change of buckling shapes when crossing this length delimitation. Focusing on the shape of the first buckling mode, the position with the maximum deflection locates at the middle of the beam when l_{SMA} is below 18.5 μm. Because the constraint from the springs weakens as l_{SMA} increases, the deflection at the free-end is increasing and the part of the beam, which is close to its fixed end, becomes flat gradually. The shapes of two modes also become similar progressively, which explains the narrowing down of the gap of the first and second critical forces. After crossing the delimitation ($l_{\mathrm{SMA}} > 18.5$ μm), the deflection at the free-end of the first mode continues to increase while it decreases for the second mode. Therefore, the critical forces also separate from each other.

Overall, to satisfy Eq. 6.12, a larger difference between these two critical forces is preferable, because it gives larger tunable range of the compressive force for buckling. Therefore, it is advisable that l_{SMA} should be carefully selected to avoid falling into the proximity of the value that the change of buckling shape takes place. Choosing either small or large value of l_{SMA} depends on the required buckling shape and the compressive force that the structure can provide to create buckling, which should be larger than the first-mode critical force.

Length of Beam I l_{Si}: Changing the length of Beam I, l_{Si}, shows a similar trend to that of l_{SMA}, as depicted in Fig. 6.9. Two critical forces decrease considerably with increasing l_{Si} and there also exists a length delimitation of 24.5 μm, for which the difference between the two critical forces reaches the smallest value. The reason can also be contributed to the change of buckling shape by observing the shape changes at different l_{Si} listed in Table 6.5. A subtle distinction from Fig. 6.9 is that there are local extrema in the curves: in

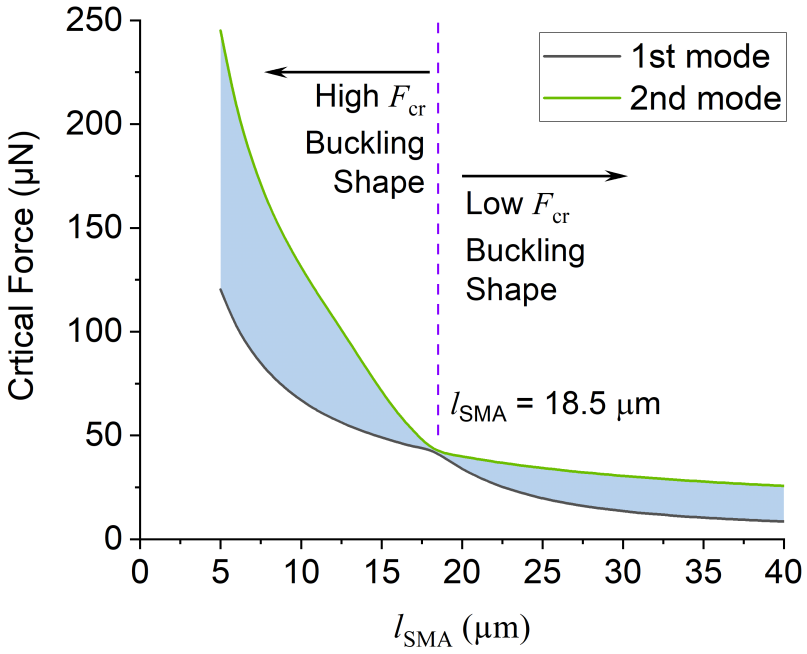


Figure 6.8.: Influence of the length of Beam II, l_{SMA}, on the first- and second-mode critical force of buckling. Shaded area represents the region that the compressive force should fall into to create buckling.

Table 6.4.: Shapes of the first and second buckling modes of Beam I at different lengths of l_{SMA}. The left ends are fixed in all scenarios drawn.

l_{SMA}	First Buckling Mode	Second Buckling Mode
10 μm		
18.5 μm		
30 μm		

the length range below the delimitation value, the curve of the first-mode critical force of buckling has a local minimum at l_{Si} = 16 μm; meanwhile, a local maximum of the second-mode critical force is witnessed at l_{Si} = 32.5 μm, which is in the length range above the delimitation value. To conclude, l_{Si} should be chosen carefully likewise to prevent from falling into the proximity of the length, for which the change of buckling shape takes place. Because the difference between the two modes of critical force at the local minimum (l_{Si} = 16 μm) is larger than the the difference at local maximum (l_{Si} = 32.5 μm), and the first-mode critical forces at the two local extrema are comparable, the length value at the local minimum is the optimal value to choose.

Width of Beam I w_{Si}: The third parameter investigated is the width of of Beam I, w_{Si}. Its relationship with the critical forces of buckling shows a different trend, as drawn in Fig. 6.10. The second-mode critical force increases nearly linearly with increasing

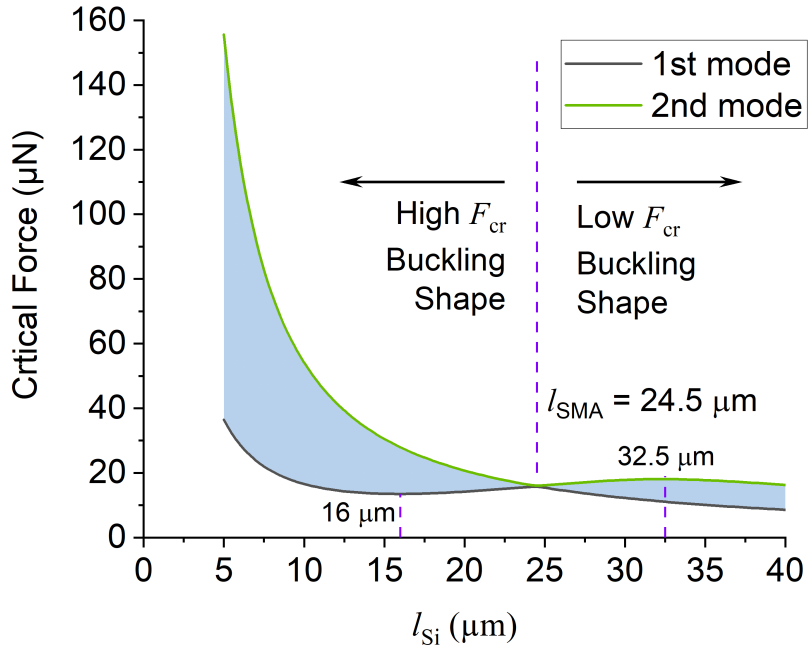


Figure 6.9.: Influence of the length of Beam I, l_{Si}, on the first- and second-mode critical force of buckling. Shaded area represents the region that the compressive force should fall into to create buckling.

Table 6.5.: Shapes of the first and second buckling modes of Beam I at different lengths of l_{Si}. The left ends are fixed in all scenarios drawn.

l_{Si}	First Buckling Mode	Second Buckling Mode
10 μm		
16 μm		
24.5 μm		
32.5 μm		

w_{Si}. Meanwhile, variation of the first-mode critical force is comparably trivial in the investigated width range. Besides, the curve of first-mode critical force also shows a local minimum at w_{Si} = 0.7 μm. To conclude, the first-mode critical force shows less sensitivity to the width of Beam I. A larger width is favorable for enlarging the difference between the first- and second-mode critical force. This can also increase the etching time from the view of fabrication, which leads to unwanted over-etching phenomenon. A compromise between these two aspects gives the appropriate range of w_{Si}, whereby the local minimum of w_{Si} is the optimal value to choose.

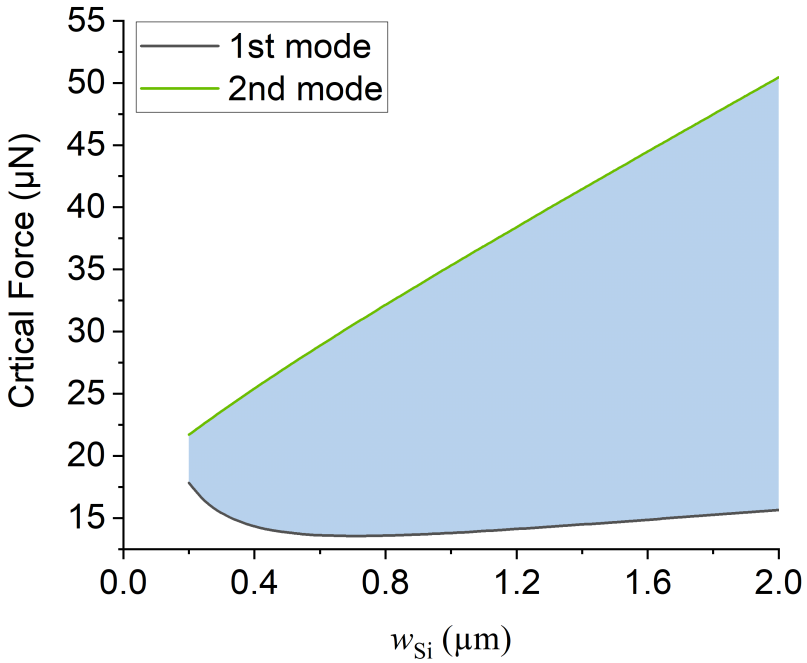


Figure 6.10.: Influence of the width of Beam I, w_{Si}, on the first- and second-mode critical force of buckling. The shaded area represents the region that the compressive force should fall into to create buckling.

6.3.3. Compressive Force and Thermal Moment

As discussed previously, the compressive force F_{eq} for creating buckling and the thermal moment M_{eq} for state-switching of Beam I originate from the residual strain and thermal strain in SMA layer of Beam II and they are defined by Eqs. 4.15 and 4.17, respectively. Parameters relating to material properties in these two equations, e.g., E_{SMA} and residual stress in SMA layer, cannot be tuned during structure fabrication, therefore these parameters are not discussed here. Modifiable geometrical parameters include the SMA layer thickness h_{SMA} and the width of the two beams (w_{Si} and w_{SMA}). Among them, the widths have a linear relation to both F_{eq} and M_{eq}, and the relation between h_{SMA} and F_{eq} is also linear. Hence, only the interaction between h_{SMA} and M_{eq} is discussed in the following.

Based on the material properties of the SMA listed in Table 5.2 and the parameters from Table 6.2, the variation of M_{eq} with increasing h_{SMA} is plotted in Fig. 6.11. Though there is a peak in the thermal moment, the corresponding thickness of the SMA is too high to be compatible with the lift-off based process, which requires the resist to be 2–3 times thicker than the deposited film. Thick resist usually poses challenges during resist coating and lithography. In the range below 1 μm, the moment increases nearly linearly with the thickness of SMA. Therefore, the major restriction on the selection of h_{SMA} comes from the limitation on the fabrication process.

Besides, one can also notice that the thermal moment is far lower than the residual moment (around 9 μN·μm). A potential solution is to introduce a low stress layer, e.g., a SiO_2 layer [50], on top of SMA layer to offset the residual moment.

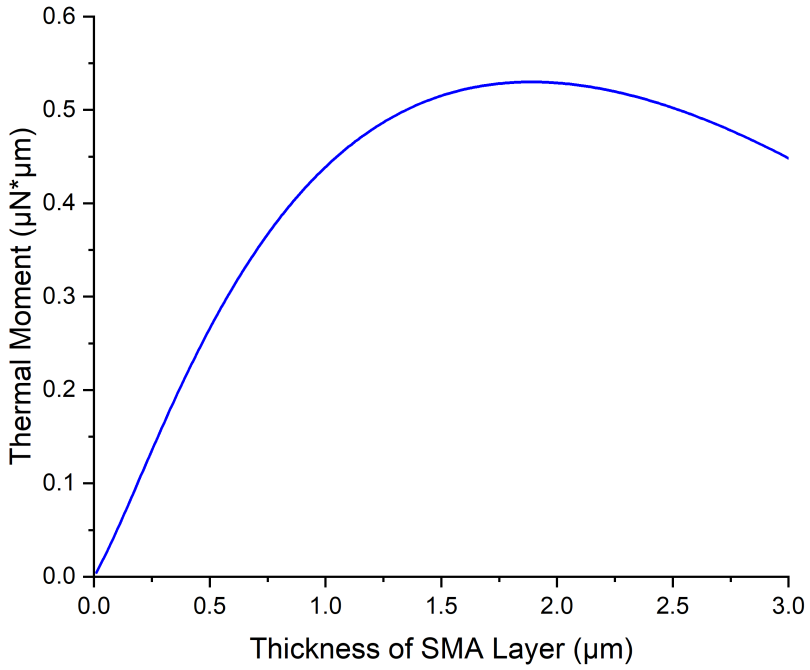


Figure 6.11.: Analytically calculated relation between the thermal moment generated by Beam II in the temperature range from room temperature to A_s and the thickness of SMA layer.

6.3.4. Bistable Stroke and Threshold Moment

In the previous sections, the parametric studies are focused on the parameters that have direct relations to geometrical variables. The parameters that describe the bistable performance, i.e., the bistable stroke Δ and threshold moment M_{th}, are influenced by the parameters investigated above and their relations are studied in the following.

Fig. 6.12 depicts the reaction moment-displacement curve of Beam I under different compressive forces, which are simulated in Comsol. Dimensions and the spring stiffnesses are based on the values in Table 6.2 and Table 6.3. The second mode critical force of buckling is 2.25 times the first mode force, hence all compressive forces studied are within the range of Eq. 6.12. Evidently, larger compressive force leads to larger bistable stroke, while higher switching moment is also required. It is worth noting that due to the limitation of the classic buckling theory, a compressive force being slightly higher than the critical force may not be able to successfully trigger the buckling of the beam. Thus the curve of $1.02F_{\text{cr}}^{1st}$ does not show bistability.

In conclusion, due to the complexity of the interplays among all the parameters discussed above, it is challenging to describe the designing process of the SMA-based bistable structure by a clear linear flow chart. While in general it can be summarized as follows: Based on the material properties of the SMA layer, the first step is to decide the length of Beam II (l_{SMA}) and have appropriate spring stiffness values (k_x, k_y and α_r). Meanwhile, according to the compressive force and thermal moment that the SMA film can provide, the geometrical dimensions of Beam I can be decided by a parametric sweep so that it

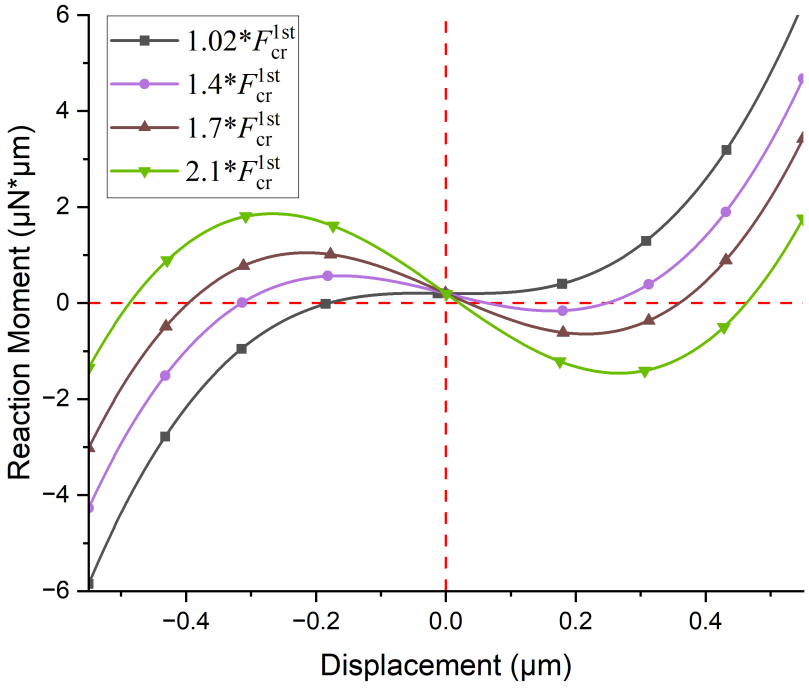


Figure 6.12.: Simulated reaction moment-displacement curves of Beam I under different compressive forces.

has appropriate critical forces of buckling. Several iterations of this process are necessary until finding the optimal design space of parameters.

6.4. Fabrication Process

Fig. 6.13 demonstrates the fabrication process of the out-of-plane bistable SMA nanoactuator. The process starts with a SOI chip and corresponding thicknesses of device layer and box layer are of 220 nm and 3 µm, respectively (Fig. 6.13(a)). Standard cleaning procedures are carried out beforehand and gold markers are patterned on the chips by lift-off prior to structure fabrication due to the demand of more than one exposure step, as introduced in Section 2.2.1.

A negative resist is firstly coated on the chips and exposed by EBL to define the patterns of nanoactuators and conductive lines. The patterns are transferred to the Si layer by RIE using the cryo process, as shown in Fig. 6.13(b). Etching gases for this process include SF_6, O_2 and Ar. The structures are then released in HF-based solution. Next, the chips are transferred to critical point drying to avoid potential damage on free-standing structures caused by liquid stiction (Fig. 6.13(c)). Afterwards, a bilayer PMMA is coated and exposed to define the sputtering window for the SMA layer, which is added to the chip by DC magnetron sputtering. Followed by a lift-off process, the SMA layer remains only on the two beams in the middle (Fig. 6.13(d)). Lastly, the chips are annealed to crystallize the SMA layer.

Moreover, it is important to note that the bilayer PMMA for lift-off in Step (d) needs to be thick enough to fill the height difference between the free-standing structures and substrate after releasing. For instance, if this height gap is 1 µm after step (c) and the thickness of the SMA layer is 300 nm, the bilayer PMMA needs to be at least 1.6 µm to fulfill a successful lift-off. The upper limit of the SMA layer thickness that can be deposited is therefore restricted by the thickness of the bilayer PMMA that can be coated.

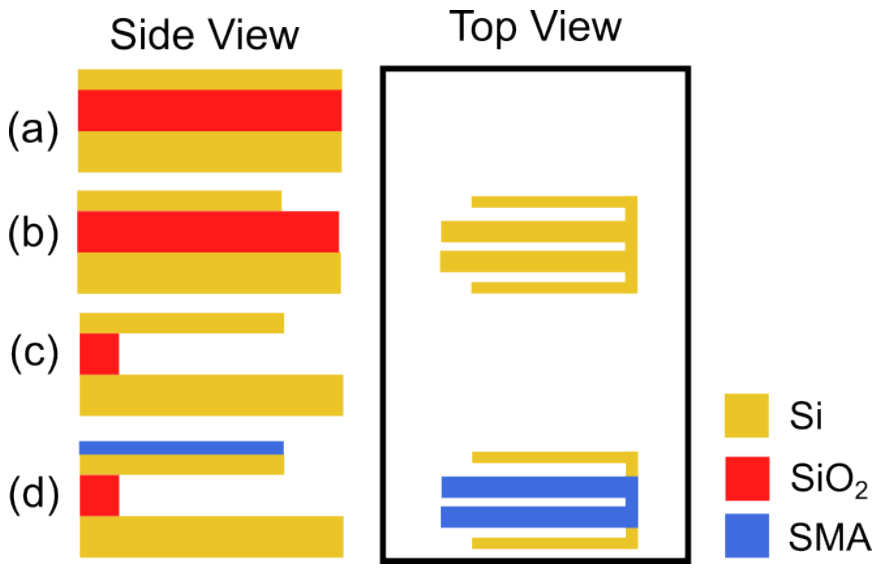


Figure 6.13.: Fabrication processes of the SMA nanoactuator with out-of-plane bistability. Starting with (a) the SOI chip, the patterns of nanoactuators and conductive lines are defined by EBL and transferred to the device layer by RIE in step (b). Afterwards, the structures are released by a wet-etching step (c). Lastly, the SMA layer is locally deposited on the two middle beams by lift-off process and annealed in step (d).

6.5. Characterization of the Fabricated Nanoactuators

Beyond the expectation, after following the fabrication process in Fig. 6.13, it is found that there is unforeseen and significant internal stress in the Si layer of the fabricated devices, which greatly influences the functionality of the original design. In the following, this internal stress and its influences on the performance of bistable nanoactuators will be investigated.

6.5.1. The Pure Si Structures

After releasing the patterned Si layer and before depositing the SMA layer (Step (c) in Fig. 6.13), it is observed that there is large deflection of the free-standing beams. Fig. 6.14 demonstrates two of these bent beam structures, which have the deflections of around 4.1 µm and 4.8 µm , respectively. The structural dimensions are

- in Fig. 6.14(a): $l_{SMA} = 30$ µm, $w_{SMA} = 2.2$ µm, $l_{Si} = 25$ µm, $w_{Si} = 1.0$ µm,

- in Fig. 6.14(b): l_{SMA} = 40 µm, w_{SMA} = 2.7 µm, l_{Si} = 25 µm, w_{Si} = 0.7 µm,

and the thickness of both structures is h_{Si} = 220 nm. Besides, the curvatures of Beam II and Beam III vary along the length direction, which are marked out by the dashed frames in Fig. 6.14. In the light-green frame regions, Beam II and Beam III have larger curvatures than Beam I and Beam IV, as the sidewall of Beam IV is clearly seen. While in the yellow frame regions, Beam I and Beam IV are "flatter" than Beam I and Beam IV indicating smaller curvatures.

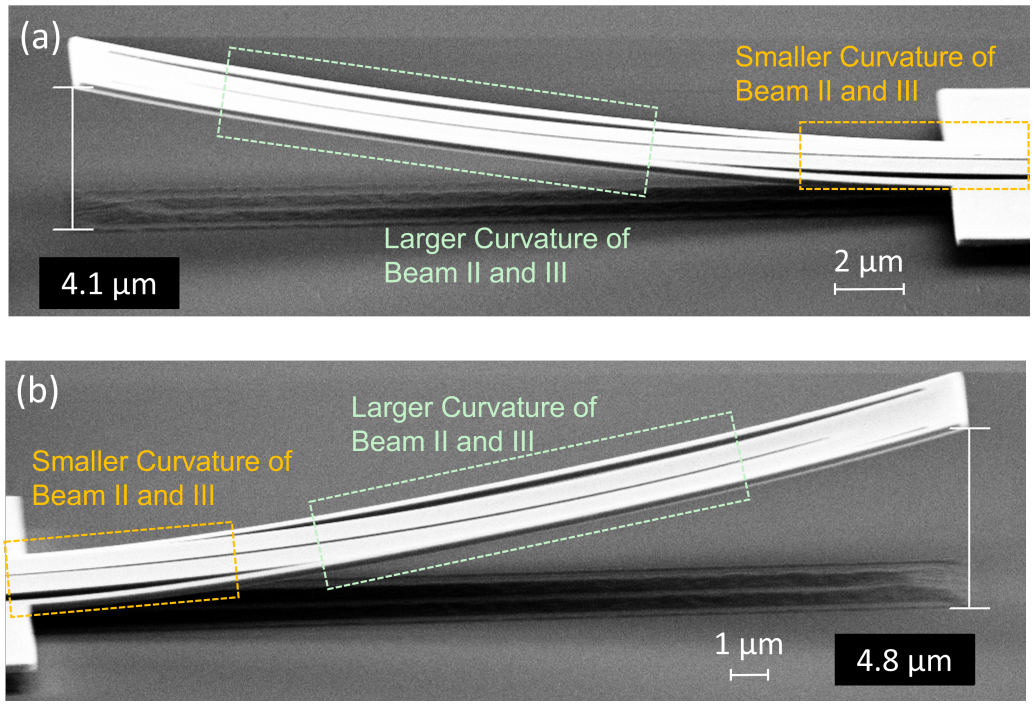


Figure 6.14.: SEM images of two four-beam structures of pure Si after Step (c) of the fabrication process. Structural dimensions are: (a) l_{SMA} = 30 µm, w_{SMA} = 2.2 µm, l_{Si} = 25 µm, w_{Si} = 1.0 µm; (b) l_{SMA} = 40 µm, w_{SMA} = 2.7 µm, l_{Si} = 25 µm, w_{Si} = 0.7 µm. The dashed frames indicate the regions with curvatures difference between Beam I and Beam II. The deflections at the free-ends of the beams are marked out in the pictures.

Table 6.6 lists the deflection values of additional four-beam structures of pure Si with different dimensions. By comparison among the values, the influences of structural dimensions on the deflection can be qualitatively summarized as:

- increasing l_{SMA} has negligible influence on deflection,
- increasing w_{SMA} leads to larger deflection,
- increasing l_{Si} leads to larger deflection,
- increasing w_{Si} leads to minor decrease of deflection.

Another experiment is conducted to intentionally fracture the front end of the structure in Fig. 6.14(a), which connects the four beams. This is achieved by using a nanoindenter, as shown in Fig. 6.15. It can be noticed that the Beam II and Beam III in the middle remain

Table 6.6.: Deflections of the designed bistable beam structures of pure Si with different combinations of dimensions.

l_{SMA}/µm	w_{SMA}/µm	l_{Si}/µm	w_{Si}/µm	Deflection/µm
20	2.2	15	0.7	2.3
30	2.2	15	0.7	2.2
		20	0.7	3.5
		25	0.7	4.4
			1.0	4.1
40	1.8	25	0.7	3.3
	2.2	15	0.7	2.2
		25	0.7	4.2
			1.0	4.2
			1.4	4.0
	2.7	25	0.7	4.8

bending up and the deflection increases about 1 µm. On the contrary, Beam IV becomes flat. Beam I is broken during loading the nanoindenter and flies out of the scope of the current view. It can be assumed that Beam I is at the same deflection state as Beam IV due to symmetry.

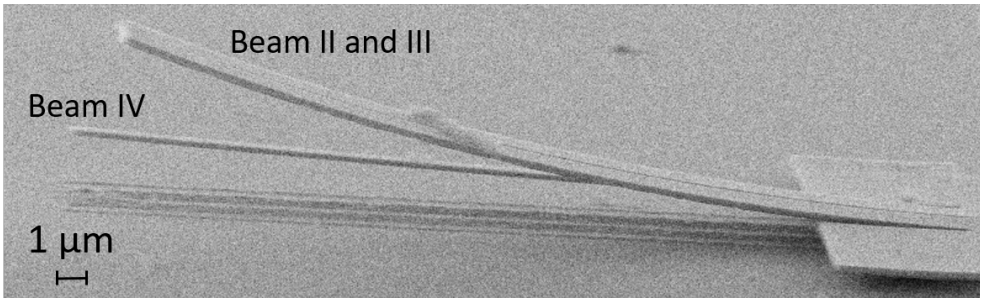


Figure 6.15.: After breaking the front connecting the four beams of the structure in Fig. 6.14(a), Beam II and Beam III are still bent up whereas Beam IV becomes flat (Beam I is broken and flies out of the scope of the view).

Based on the observations in the SEM images of Fig. 6.14 and Fig. 6.15, the following phenomenological conclusions can be drawn.

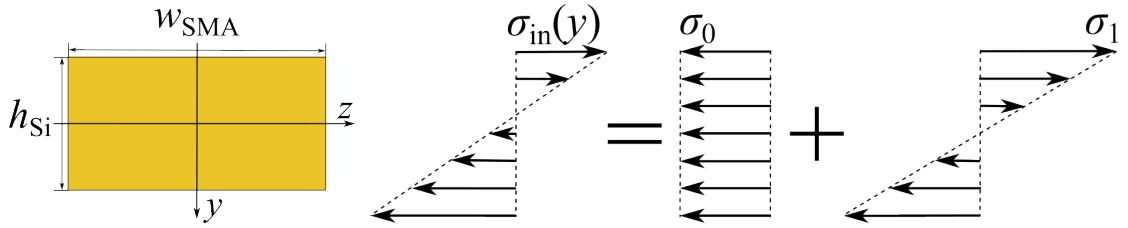
1. There is position-dependent internal stress in the Si layer that results in the deflection of the single layer structure. The stress in Beam II and Beam III is considerable, whereas it is negligible in Beam I and Beam IV.
2. The curvature variations indicate the buckling of Beam II and Beam III due to the compressive internal stress, as bending by constant moment leads to also constant curvature (Eq. 4.16).

3. The internal stress does not originate from the fabrication of the SOI wafers, which is supposed to distribute evenly across the wafer. This stress is introduced during the fabrication process (Fig. 6.13).

Based on these summarized points, by referring to the cross-section of Beam II (same for Beam III) in Fig. 6.16(a), the internal stress $\sigma_{\text{in}}(y)$ in Si layer is assumed to have a linear distribution along the thickness ($y-$direction). Accordingly, as illustrated in Fig. 6.16(b), $\sigma_{\text{in}}(y)$ is composed of two terms as

$$\sigma_{\text{in}}(y) = \sigma_0 + \sigma_1 \frac{2y}{h_{\text{Si}}}, \tag{6.20}$$

where the first uniform term σ_0 is negative and contributes to the buckling phenomenon; the second proportional term generates a bending moment. It is worth noting that $\sigma_{\text{in}}(y)$ does not need to strictly meet the distribution in Fig. 6.16(b) that the stress at $y = -h_{\text{Si}}/2$ is positive; it may also be negative if the slope of the linear function Eq. 6.20 is small.



(a) Cross-section of Beam II. (b) Decomposition of the linearly distributed internal stress $\sigma_{\text{in}}(y)$.

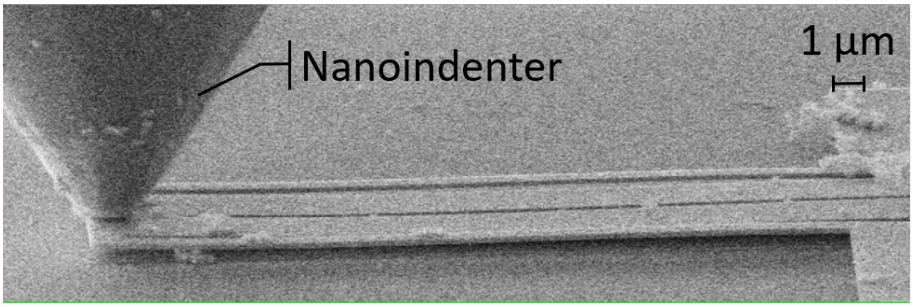
Figure 6.16.: (a) The cross-section of Beam II and (b) the linear distribution of the internal stress $\sigma_{\text{in}}(y)$ in the Si layer. $\sigma_{\text{in}}(y)$ is decomposed into a uniform term σ_0 and a proportional term.

Due to the non-linearity of buckling and the interactions between buckling and elastic bending, it is difficult to analytically determine the two unknown terms σ_0 and σ_1. In other words, both compressive stress and bending moment can lead to deflection of beam, while the proportions of contribution from these two terms cannot be exactly determined. Nevertheless, the ranges of the two terms can still be given by considering the extrema situations. If the bending is exclusively caused by buckling, the maxima of σ_0 is simulated to be 13.5 MPa; on the contrary, the bending only caused by moment gives the maxima of σ_1/h_{Si} to be 1980 MPa/µm, which is calculated based on Eqs. 4.9 and 4.16.

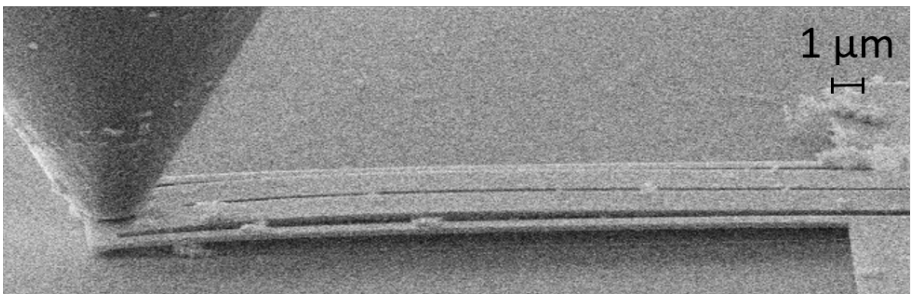
Because the fabrication results reveal the compressive stress in Beam II and Beam III, which is different from the assumptions in Section 6.2 and 6.3 that there is no internal stress in the Si layer, hereinafter the analyses are reversed that Beam II (and Beam III) is buckled on the elastic supports provided by Beam I (and Beam IV).

To verify the buckling of Beam II and Beam III, a displacement load is vertically applied at the front end of the structure with the nanoindenter. An example is demonstrated in Fig. 6.17, in which the structural dimensions are $l_{\text{SMA}} = 30$ µm, $w_{\text{SMA}} = 2.2$ µm, $l_{\text{Si}} = 25$ µm, $w_{\text{Si}} = 0.7$ µm, respectively. The loading rate is set at 0.05 µm/s to satisfy the quasi-static condition. Theoretically this out-of-plane load should be able to switch the beams from

one buckled state to the other. The displacement of the nanoindenter and the reaction force are recorded.



(a) Buckled down state.



(b) Buckled up state

Figure 6.17.: The snap-through of the four-beam structure upon loading at the front end by the nanoindenter. Dimensions of structure are l_{SMA} = 30 μm, w_{SMA} = 2.2 μm, l_{Si} = 25 μm, w_{Si} = 0.7 μm.

The initial shape of the structure in Fig. 6.17 is similar to that in Fig. 6.14, and with the deflection of 4.4 μm (given in Tab. 6.6). Because of the downward loading, the structure gradually becomes flat, and Beam II (and Beam III) deforms into a concave shape (Fig. 6.17(a)). With a slight increment of the loading, the middle of Beam II and Beam III quickly snap to a convex buckled-up shape (Fig. 6.17(b)). The snap-through stroke is approximately 500 nm, corresponding to 1.67% of beam length. Unfortunately, the snapped state is not able to be held. When the nanoindenter is retracted at the same rate of 0.05 μm/s, the structure firstly snaps back the to shape in Fig. 6.17(a), and restores to the initial deflected shape at the end. Explanation of this snapping-back is that the bending moment generated by the term σ_1/h_{Si} is larger than the moment required for state switching. A similar phenomenon has also been studied in [49]. Same loading cycles are applied to structures of different dimensions, and the snapping phenomenon is witnessed on all the tested structures.

The displacement-reaction force curves of two tested structures, i.e., Structure A with dimensions l_{SMA} = 20 μm, w_{SMA} = 2.2 μm, l_{Si} = 15 μm, w_{Si} = 0.7 μm and Structure B with dimensions l_{SMA} = 30 μm, w_{SMA} = 2.2 μm, l_{Si} = 25 μm, w_{Si} = 0.7 μm, are plotted in Fig. 6.18. The negative force represents the downward direction. Because the non-linear snapping does not take place at the loading position, the curves show a simple linear relation before the nanoindenter touches the substrate. The slopes of the fitted lines

indicate the stiffnesses of the structures, which are 0.522 N/m for Structure A and 0.106 N/m for Structure B, respectively. In Chapter 4, Eq. 4.52 has defined the bending stiffness of cantilever beams. The analytically calculated stiffnesses by Eq. 4.52 for the two structures here are 0.437 N/m and 0.114 N/m, respectively, which are close to the fitted values.

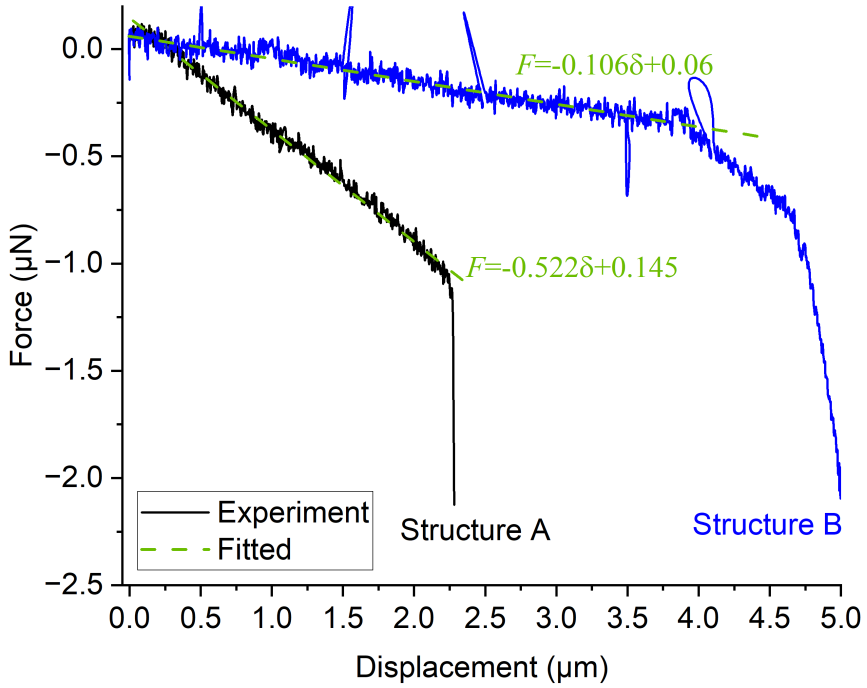


Figure 6.18.: Relations between the displacement and the reaction force (solid lines) recorded by the nanoindenter, which can be fitted by linear functions (dashed lines). Dimensions of the tested structures are: l_{SMA} = 20 μm, w_{SMA} = 2.2 μm, l_{Si} = 15 μm, w_{Si} = 0.7 μm for Structure A; l_{SMA} = 30 μm, w_{SMA} = 2.2 μm, l_{Si} = 25 μm, w_{Si} = 0.7 μm for Structure B.

6.5.2. The SMA Nanoactuators

Continuing the fabrication process in Fig. 6.13, a 300 nm-thick $Ti_{54}Ni_{31}Cu_{15}$ layer is selectively deposited on the area of Beam II and Beam III by lift-off. Fig. 6.19 compares the initial deformation of the nanoactuators before and after the annealing of the SMA layer. The dimensions of the structures are the same as the structure in Fig. 6.14(a). Due to the change of the residual stress from negative to positive in the SMA layer by the annealing process, the structure experiences a significant shape variation from almost flat shape before annealing to largely bent shape after annealing (deflection about 8.6 μm), which is twice as large as the deflection before SMA deposition. It is worth noting that the deformation before annealing is probably hindered by the insufficient space between the beam and the SiO_2 substrate. To rephrase this, because of the negative residual stress in SMA layer, there can be noticeable downwards deflection if the gap size between the beam and the substrate is large enough. Additionally, severe sidewall deposition originating from sputtering process of $Ti_{54}Ni_{31}Cu_{15}$ film is observed, causing the skirt-like residuals along the patterns of SMA layer.

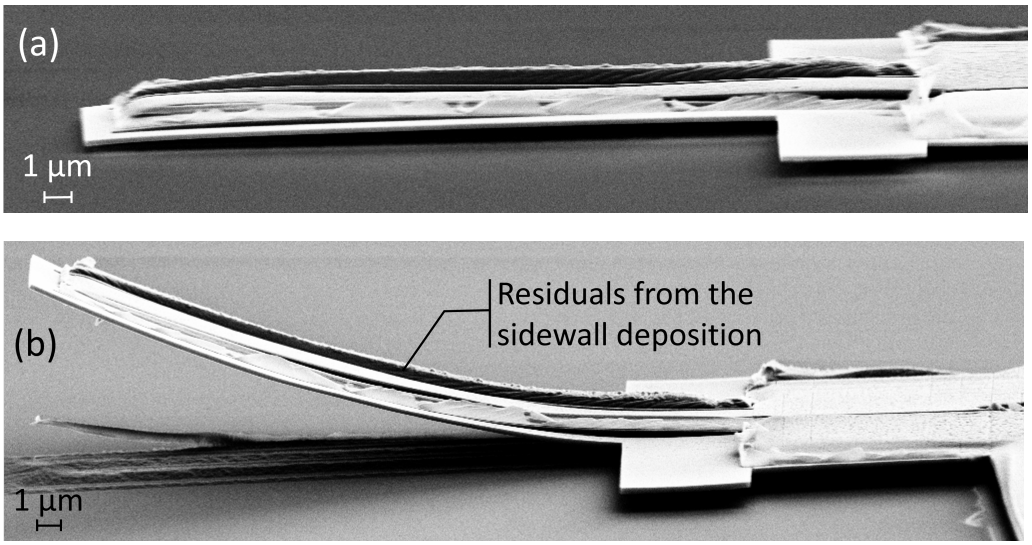


Figure 6.19.: SEM images of the structure in Fig. 6.14(a) after depositing the SMA layer (Step (d) of fabrication process), (a) before and (b) after the annealing of SMA layer.

Compression tests with the nanoindenter are also conducted with the nanoactuators after being annealed. Fig. 6.20 presents the displacement and force versus time curves of the nanoactuator with the following dimensions: $l_{SMA} = 40$ µm, $w_{SMA} = 2.2$ µm, $l_{Si} = 25$ µm, $w_{Si} = 1.4$ µm. Several turning points (Point (a)–(e)) are marked in the figure, which correspond to the SEM images showing the deformed structural shapes in Fig. 6.21. Before the deposition of the SMA layer, the structure has a deflection of about 4 µm, whereas this value increases to around 9.8 µm after the annealing of SMA, as demonstrated in Fig. 6.21(a).

Upon the quasi-static loading at the free-end of the nanoactuator, the deflection decreases linearly with the displacement until point (b) that Beam I and Beam IV begin to touch the substrate (the yellow dashed frame in Fig. 6.21(b)). These two beams possess a concave shape that is similar to a buckled-down shape. Afterwards, the reaction force surges from 10.1 µN at point (b) to 84.4 µN at point (c) with a small displacement increment from 7.7 µm to 10.0 µm. During this loading process, Beam IV gradually buckles up, which is highlighted by a yellow dashed frame in Fig. 6.21(c). This fast growth of the force can be attributed to the switching of the buckled shapes of Beam IV that additional energy input is required to overcome the energy barrier. Possibly due to the imperfections during the fabrication, the shape change of Beam I does not take place simultaneously with Beam IV despite the structural symmetry. This delay is explicitly manifested by another increase of force from point (c) to point (d) within an almost negligible displacement change. Having reached the peak value of 108.7 µN at point (d), the force witnesses a noticeable drop to 49.3 µN at point (e). As shown in Fig. 6.21(e), this process is accompanied by the continuous buckling up of both Beam I and Beam IV without displacement variation, because the instable position of the bistable buckled beams has been crossed and no further energy input is required. Though the snapping of Beam I is not completely finished, Beam IV exhibits a buckled-up shape with maximum deflection of around 900 nm, which is 3.6% of beam

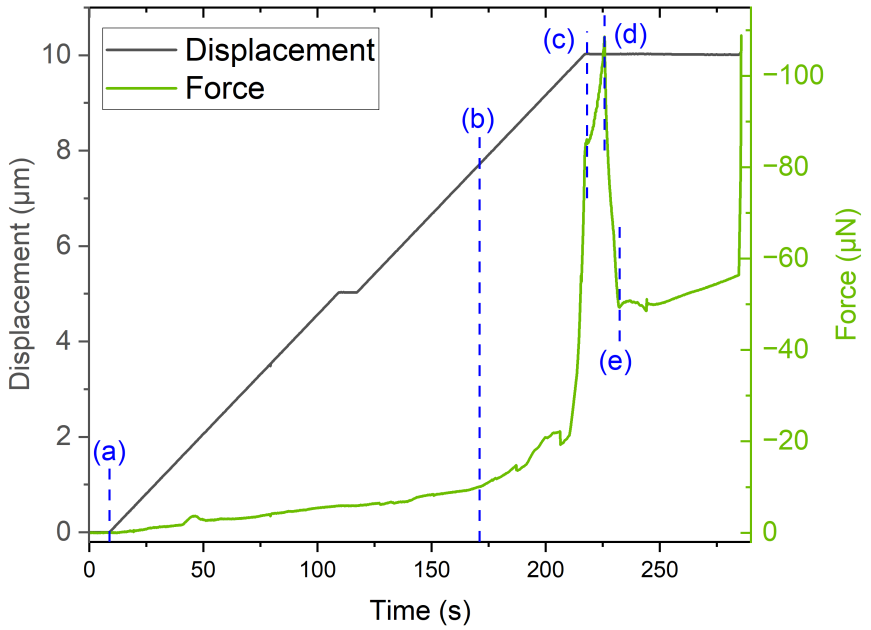


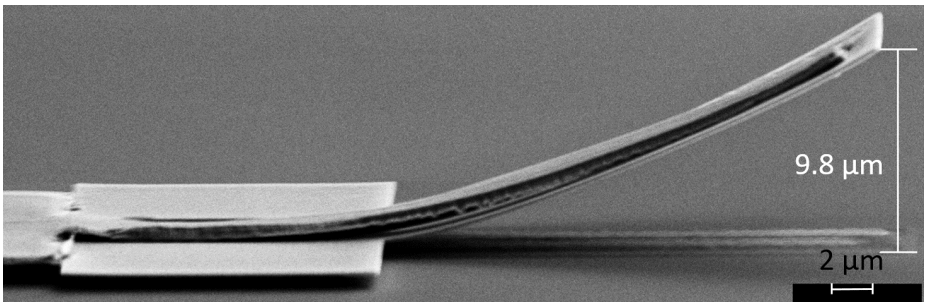
Figure 6.20.: Loading curves (Displacement and Force versus Time) of structure with the following dimensions: l_{SMA} = 40 µm, w_{SMA} = 2.2 µm, l_{Si} = 25 µm, w_{Si} = 1.4 µm. The structural shapes at Point (a)–(e) are shown in Fig. 6.21. The minus force represents the direction.

length. Further loading leads to another surge in the force as the nanoindentor touches the substrate. These buckled-up shapes are also not able to be held during unloading, and the tentative explanation is the large moment induced by the internal stresses in both SMA and Si layers.

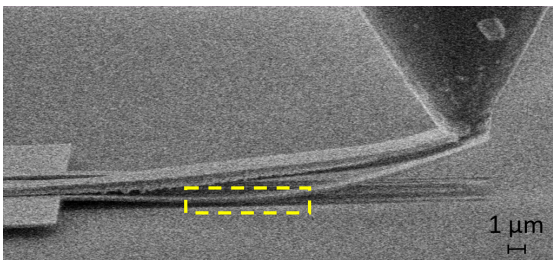
With varying geometrical dimensions, the snapping phenomenon takes place only on certain nanoactuators, which is different from the structures without SMA layer that the snapping phenomenon exists commonly. The relations between the snapping and the geometrical dimensions are still under investigation.

6.6. Discussion

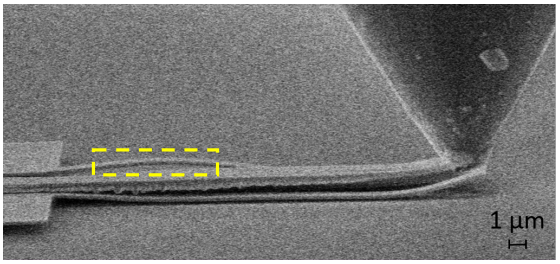
Because the bistability of the proposed design of nanoactuators is based on the buckling of beams, the analyses of buckling play a key role in the bistable performance. As presented above, in the experiments the major reason causing the deviation from the design is the internal stress introduced during the fabrication. Though this stress has a compressive component that is beneficial for buckling, its gradient component generates a moment having the same direction as the residual moment from the SMA layer. Therefore, the overall threshold moment that the state-switching moment needs to overcome is largely elevated, and in the worst situation, the bistable property brought by buckling completely disappears.



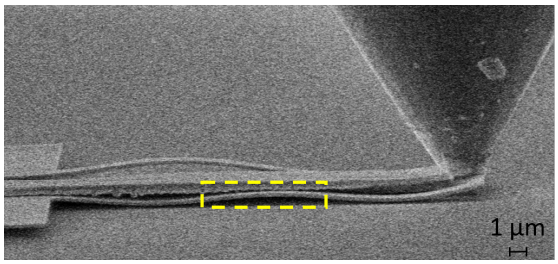
(a) Initial shape.



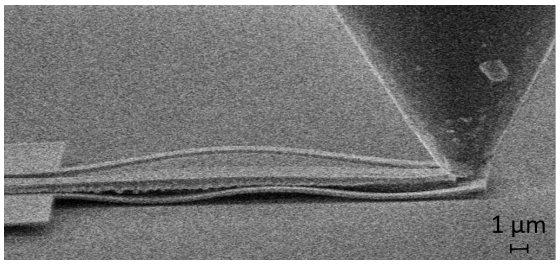
(b) Beam I and Beam IV begin to touch the substrate.



(c) Beam IV begins to buckle up.



(d) Beam I begins to buckle up.



(e) Final buckled shape.

Figure 6.21.: SEM images show the deformed structural shapes during the loading process. The corresponding moments (a)–(e) are marked in Fig. 6.20.

Another influence brought by this internal stress is the coupling effect between the compressive term for buckling and the gradient term for bending, both of which lead to deflection. As depicted in Fig. 6.22, for a cantilever beam that is already bent because of the moment M, the compressive force F for buckling also generates a bending moment to the beam, especially in the situation of large deflection. This coupling effect complicates the analyses of the internal stress with gradient. Furthermore, the moment M leads to changes of the critical buckling force F_{cr} as deformed structures are more sensitive to buckling loads. For instance, Michael and Kwok [49] have derived the change of critical buckling force of a beam in the presence of external and residual moments. This dependency poses another challenge of the analyses of buckling performance.

Consequently, understanding the origin and the generation mechanism of this position-dependent internal stress is of great importance for the amelioration of the bistable design. In-situ characterization techniques such as Raman spectroscopy [166], which reveals stress change by monitoring shift of Raman spectrum locally, complemented by

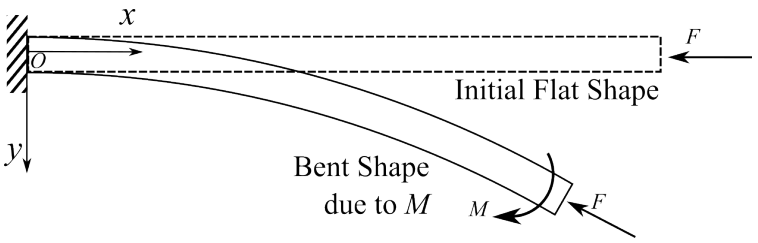


Figure 6.22.: The coupling effect between the bending moment M and the compressive force F for buckling.

sophisticated non-linear finite element simulations, are potential methods to reveal the actual stress distribution. On this basis, further improvements to the structural design should concentrate on compensation of the undesirable moments induced by the residual stresses. Feasible approaches include precise stress tuning in the SMA layer and adding a third stress compensation layer so that Eq. 4.17 can equal to zero in the ideal situation.

6.7. Summary

In this chapter, a bistable nanoactuator design composed of four cantilever beams is proposed by exploiting the residual stress in the SMA layer to introduce the buckling of beams. The bistable mechanism is based on that the buckled beams naturally possess two stable states. The state-switching load comes from the thermal moment and the moment induced by the shape memory effect of SMA layer. The buckled beam is analyzed as a cantilever beam supported by springs with finite stiffnesses, and the influences of geometrical parameters on the critical buckling force are investigated on this basis. After structural fabrication, it is found that there is noticeable internal stress in the Si layer resulting in significant moment, which hinders the bistable performance. By mechanically applying a force at the free-end of the nanoactuator, snapping phenomena are witnessed with a stroke about 3.6% of the beam length.

7. Integration of SMA Nanoactuator with Directional Coupler

Directional coupler is one of the fundamental components in the design of photonic circuits. Its coupling performance can be tuned by controlling the gap between the waveguides. The SMA/Si bilayer nanoactuator introduced in Chapter 5 is a potential candidate to perform this mechanical tuning. In this chapter, the multiphysical structure, i.e., the nanoactuator integrated directional coupler, is designed, fabricated and characterized. The photonic, mechanical and electric parts are designed separately. Parametric studies are performed to ensure the optimal performance of the key component, the directional coupler. A fabrication process involving three exposure steps is proposed together with the fabrication results. The performance of the photonic circuits is characterized by transmission tests. Lastly, the revealed fabrication problems and potential improvements to this structure are discussed.

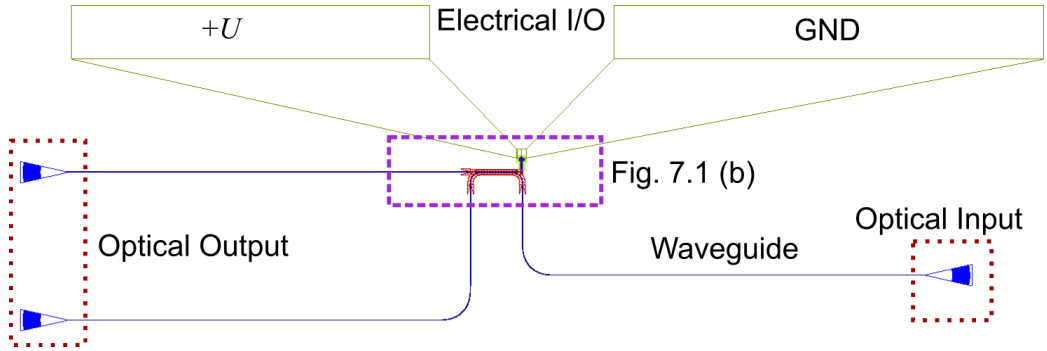
7.1. Design

It has been introduced in Chapter 3 that light propagation can be coupled from a bus waveguide into a neighboring receiving waveguide with the help of a directional coupler. The coupling efficiency is dependent on the gap between the two waveguides. The idea behind the design is that by actively controlling the waveguide distance with the help of the nanoactuator, the coupling efficiency is tuned by the nanoactuator stroke accordingly. Therefore, the nanoactuator functions as a light switch in controlling the propagation path. A general sketch of the device is depicted in Fig. 7.1, which can be decomposed into several different parts: Optical input/output (I/O), waveguides, directional coupler, nanoactuator, electric circuit and I/O. Designs of these parts are introduced in the following in detail.

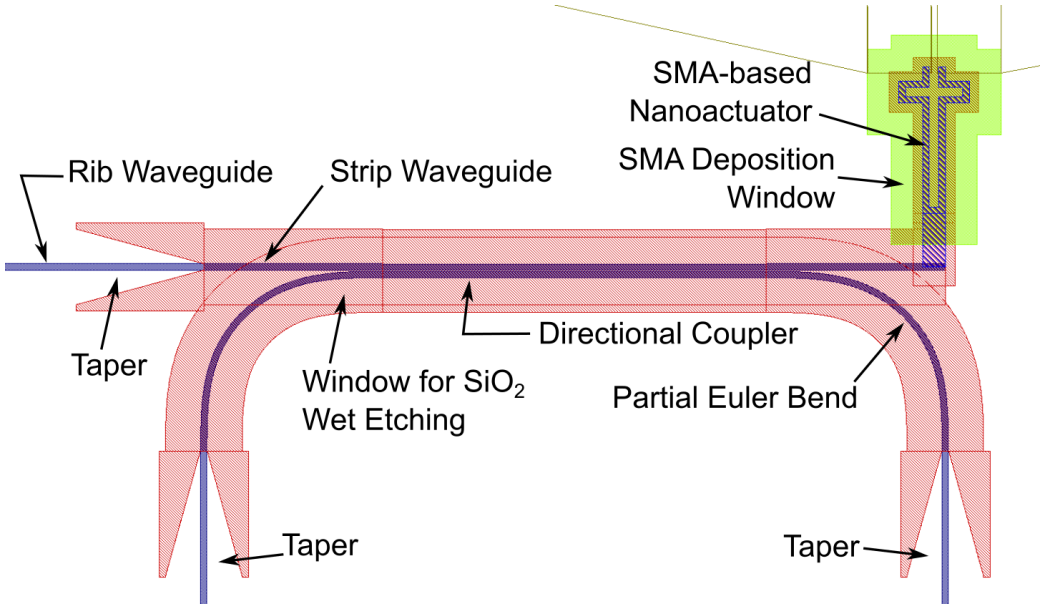
7.1.1. Optical I/O

As introduced in Section 3.4, the grating couplers are used to match the large mode difference between the optical fiber and on-chip sub-micrometer waveguides. A focusing shallow-etching grating coupler is adopted from [130] with the following advantages:

- Allowing for air as the top-cladding. Thus, this grating coupler is compatible with the movable structures on chip where the common top cladding material, for instance SiO_2, cannot be applied.



(a) Overall View

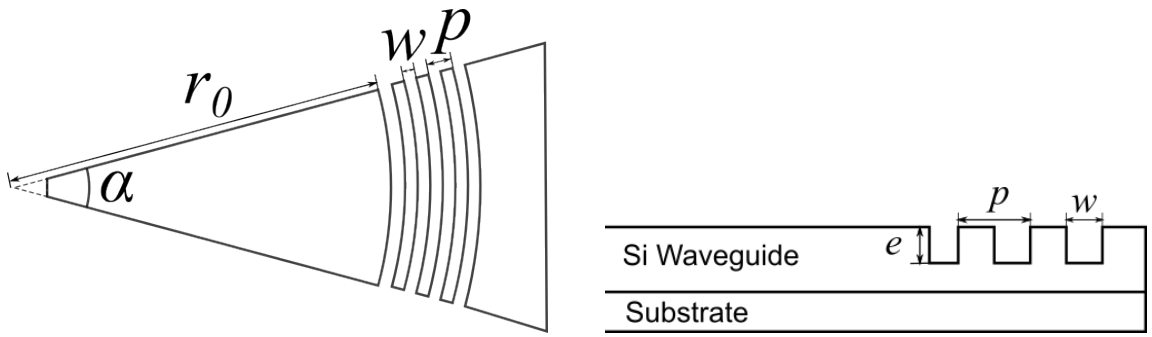


(b) Detailed View (The purple dashed frame region in Fig. 7.1(a))

Figure 7.1.: The (a) overall and (b) detailed views of the designed directional coupler integrated with SMA-based nanoactuator.

- Designing for the TE-mode light propagation, which has higher effective refractive index and converges more electromagnetic energy inside the waveguide.
- Simplifying the fabrication process as no protection layer is needed when releasing the movable structures by wet etching. Further explanations can be found in Section 7.3.
- The focusing design possesses a much smaller footprint by avoiding long taper structure between grating couplers (micrometer) and waveguides (sub-micrometer).

Fig. 7.2 shows the sketches of this focusing grating coupler with important design parameters. Exact values of these parameters are listed in Table 7.1.



(a) Top view (b) Side view

Figure 7.2.: (a) Top and (b) side view of the grating coupler with the design parameters.

Table 7.1.: Values of the design parameters of the grating coupler adopted from [130].

Design Parameters	Value
Initial radius r_0	31 µm
Focusing angle α	25°
Period p	633 nm
Groove width w	300 nm
Etching depth e	70 nm

7.1.2. Waveguides and Taper

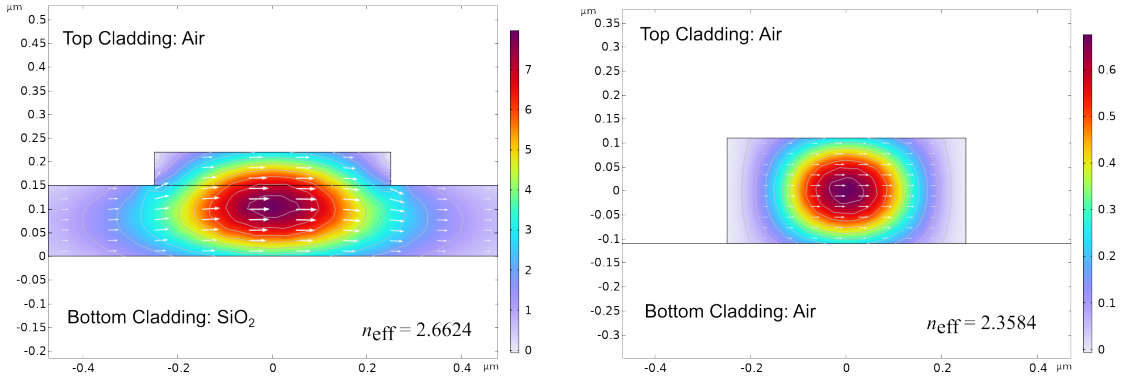
Taking into account that the grating coupler is shallow-etched while the directional coupler is deep etched in order to fabricate movable free-standing structure, two types of waveguides are needed to connect them, i.e., the rib waveguide and the strip waveguide. Design details are described in the following.

Geometrical dimensions: The (total) height for both types of waveguides is 220 nm, which is determined by the thickness of the device layer of the SOI wafer used for device fabrication. To avoid the existence of higher modes, the width of the waveguide is selected to be 500 nm. An additional parameter for the rib waveguide, the shallow etching depth, is kept to be 70 nm, which is the same as for the grating coupler described in Table 7.1.

Effective refractive index: Materials neighboring to the waveguide have a significant influence on the effective refractive indices. For the structure studies here, the top cladding for both types of waveguides is air; however, the bottom cladding for the rib waveguide is SiO_2 and for the strip waveguide is air, as the latter needs to be free-standing for functioning purposes.

Fig. 7.3 compares the contour of magnitude of the Poynting vector in the cross-section of the rib and strip waveguide, which is simulated by Comsol. Arrows in the figure show

the direction of electric field $\vec{E}$ and their lengths represent the magnitude $|\vec{E}|$. It can be learned from this figure that electromagnetic energy fluxes are well constrained in the waveguides of both types; the electric fields are almost in horizontal direction proving that the mode is very close to the theoretical TE-mode. Effective refractive indices for the rib and strip waveguides are 2.6624 and 2.3584 respectively. Therefore, it is necessary to further introduce a taper structure to connect these two types of waveguides to avoid loss due to the mode mismatch.



(a) Rib Waveguide (b) Strip Waveguide

Figure 7.3.: Contour of magnitude of the Poynting vector of (a) rib and (b) strip waveguide in TE-mode.

Taper design: To design the taper structure with appropriate dimensions, Fig. 7.4 demonstrates the relation between the slab width and the effective refractive index of a rib waveguide. It should be noted that different from the simulation in Fig. 7.3(a), both top and bottom claddings are set to be air aligning with the real situation in fabrication. The effective refractive index saturates when the slab width increases above 2 µm, therefore, the dimensions of the taper structure are determined to have the size of 2×10 µm^2 with a slab width ranging from 500 nm to 2 µm, which lowers the mode mismatch loss and does not bring additional fabrication difficulty at the same time.

7.1.3. Directional Coupler

As introduced in Section 3.3, the coupling length of a directional coupler is determined by the symmetric and asymmetric effective refractive indices with Eq. 3.26. An example is given by the free-standing directional coupler composed of two strip waveguides, which are separated by 110 nm, and both top and bottom claddings are air. The simulated symmetric and asymmetric TE-modes at wavelength 1550 nm are presented in Fig. 7.5 where the color contour represents the value of electric field in $x-$direction $\vec{E}_x$. The corresponding coupling length L_c is 36.6 µm. Fig. 7.6 shows the relation between the gap size g, the symmetric and asymmetric effective indices n_s, n_a and the coupling length L_c. Due to limitations of fabrication techniques, the gap can be well patterned when its size is no less than 100 nm, meanwhile, a larger gap size also leads to also a longer coupling length

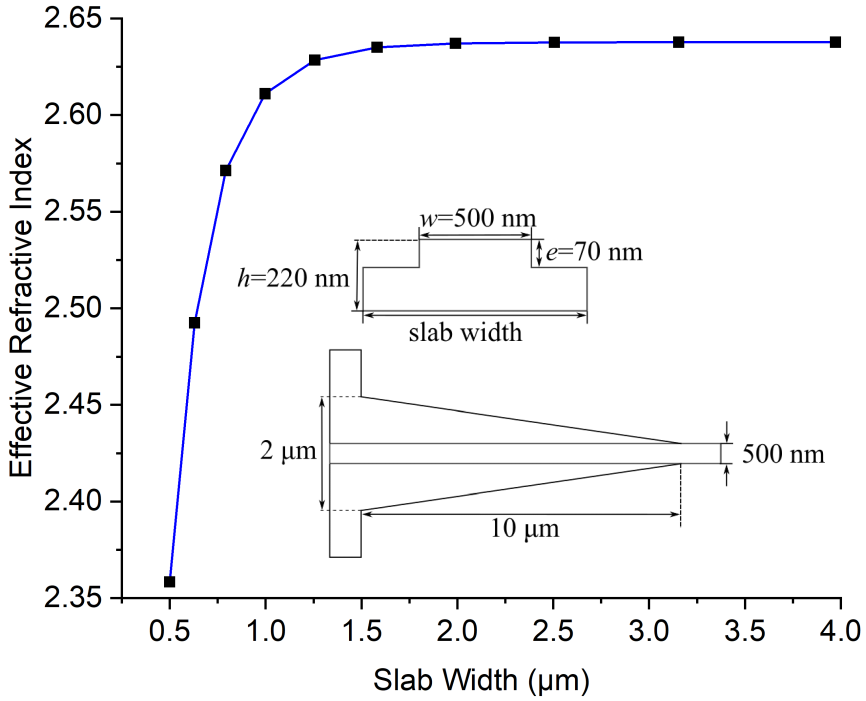
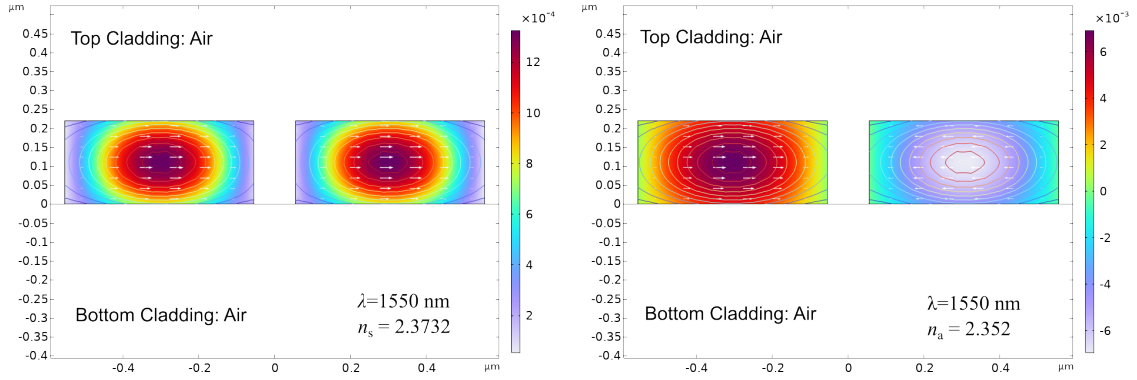


Figure 7.4.: Relation between slab width and effective refractive index of rib waveguide.

bringing further instability to the free-standing structure. Therefore, the gap size is chosen to be 110 nm for structure fabrication. Fig. 7.7 further demonstrates the oscillation of the scale of the Poynting vector between the bus waveguide and the receiving waveguide. The distance between two peaks in bus waveguide and receiving waveguide represents the coupling length L_c. Because the length of the simulated structure equals to $2L_c$, it can be clearly seen that the power flow switches back to bus waveguide after propagating this distance.



(a) Symmetric Mode (b) Asymmetric Mode

Figure 7.5.: (a) Symmetric and (b) asymmetric TE-modes of the directional coupler composed by two strip waveguides with a gap of 110 nm. Both top and bottom claddings are air. The structure is simulated at the wavelength of 1550 nm.

Additionally, the bent waveguides are adopted to connect to one waveguide of the directional coupler to separate coupling. As introduced in Section 3.2.3, the bent waveguides are in the shape of partial Euler bend, the effective radius R_{eff} is chosen to be 10 µm and the ratio $p = 0.5$.

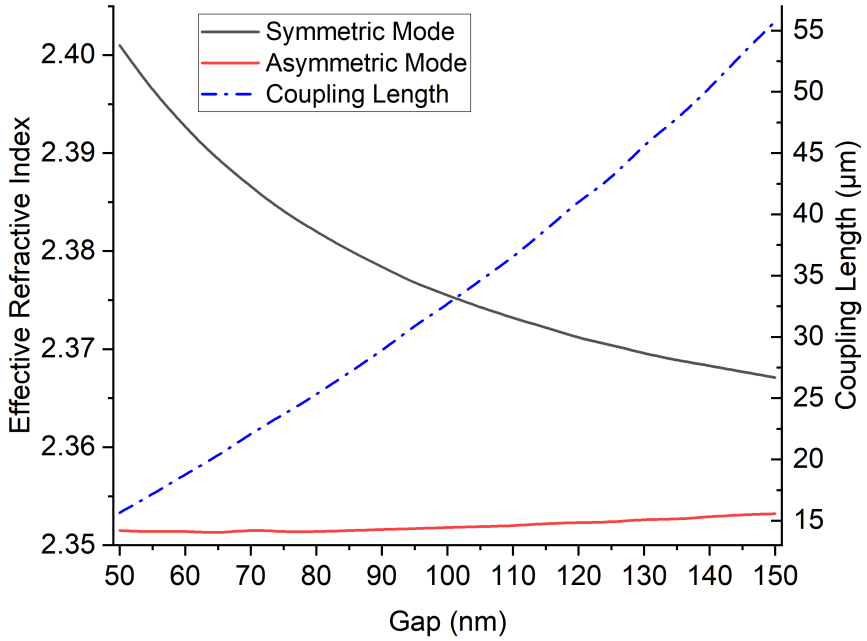


Figure 7.6.: Influence of gap size of the directional coupler on symmetric and asymmetric effective refractive indices and the coupling length simulated at the wavelength λ of 1550 nm.

7.1.4. Nanoactuator

The general considerations of designing the SMA/Si bilayer nanoactuators have been introduced in Chapter 5, which remain valid for the applications in the current chapter. However, the connected waveguide can bring additional constraint to the deflection of the nanoactuator, and this influence is analyzed in the following.

As plotted in Fig. 7.8, the length and the bending stiffness (defined by Eq. 4.13) of the nanoactuator are l_{beam} and $(EI)_{\text{eq}}$ respectively. When there is no constraint from the waveguide, the deflection of the nanoactuator induced by the equivalent bending moment M_{eq} (defined by Eq. 4.17) is δ_{eq}, which can be obtained by Eq. 4.9 and Eq. 4.16. For the waveguide structure, because it is fixed at the end connecting to the taper, it can be also seen as a cantilever beam with the length l_{wg} and bending stiffness $(EI)_{\text{wg}}$.

Assuming that the interaction between the nanoactuator and the waveguide is represented by a shear force F_{wg} in $y-$direction at the free-end of the nanoactuator, the deflection of the waveguide induced by this force is

$$\delta_{\text{wg}} = \frac{F_{\text{wg}} l_{\text{wg}}^3}{3(EI)_{\text{wg}}}. \tag{7.1}$$

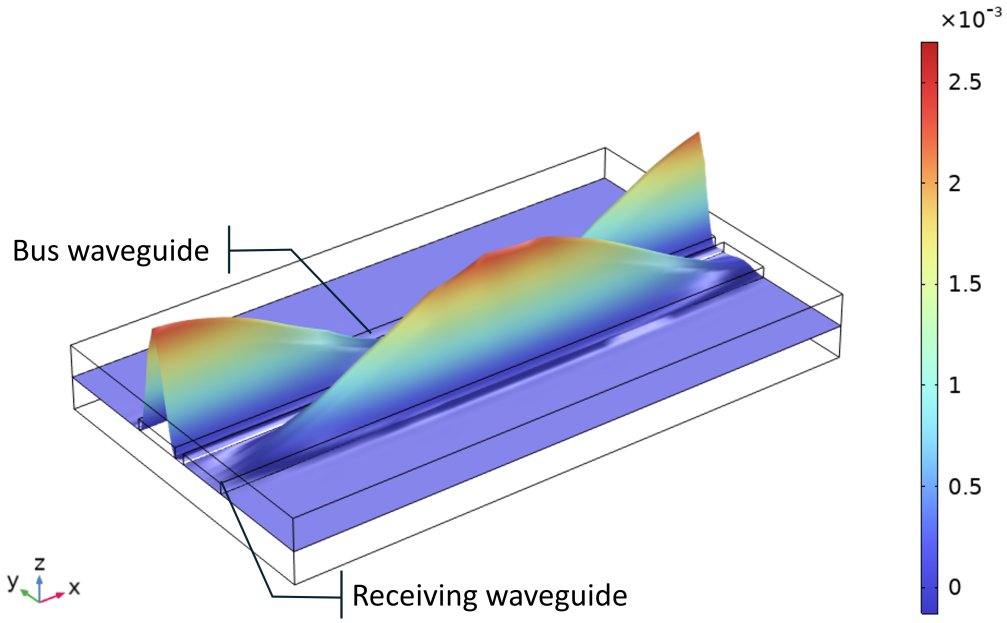


Figure 7.7.: Simulated oscillation of the scale of the Poynting vector between the bus waveguide and the receiving waveguide aroused by the imposition of symmetric and asymmetric modes. The gap of the simulated directional coupler is 110 nm and the length is twice of the coupling length (73.1 µm).

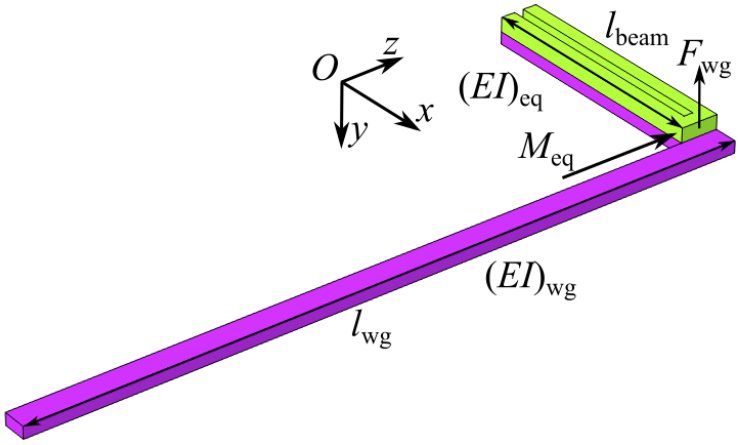


Figure 7.8.: Force analysis of the SMA/Si bilayer nanoactuator connected to a cantilever-shape waveguide.

Similarly, this force also induces a deflection on the nanoactuator

$$\delta_F = -\frac{F_{wg} l_{beam}^3}{3(EI)_{beam}}. \tag{7.2}$$

The resulting deflection of the nanoactuator needs to be equal to the deflection of the waveguide, which gives

$$\delta_{eq} + \delta_F = \delta_{wg}. \tag{7.3}$$

Combining Eq. 7.1–7.3, the ratio between δ_{eq} and δ_{wg} is

$$\frac{\delta_{eq}}{\delta_{wg}} = 1 + \frac{(EI)_{wg}}{(EI)_{beam}} \frac{l_{beam}^3}{l_{wg}^3}. \tag{7.4}$$

Since $(EI)_{wg} < (EI)_{beam}$ and $l_{beam} < l_{wg}$ in the current design, it can be concluded that the influence of the waveguide on the deflection of the nanoactuator is minor.

In addition, the nanoactuator needs to be located at a position where the interference coming from the metal layer of nanoactuator on the light propagation is negligible. This distance is set to be 2.5 μm in the layout design.

7.1.5. Electric circuit and I/O

Apart from the photonic circuit, the appropriate electric circuit and I/O is also necessary to be integrated on the chip for Joule-heating of SMA-based nanoactuators. General considerations include that the gap between the electric and photonic circuit needs to be large enough to avoid possible crosstalk. Positions of electric I/O are located near the boundaries of the chip for convenient connections to external electrical wires. In the current design, a minimum distance of 100 μm is guaranteed between electric lines and photonic waveguides, and the pads with size of 600×600 μm^2 are arranged near the edges of the chip for the wire-bonding.

7.2. Parametric Study

The performance of the nanoactuators has been thoroughly investigated in Chapter 5, therefore, the focus of this section is the parameters that have influence on the performance of the directional coupler, including the dimensions of the waveguides, the gap size and the wavelength of the propagating light. The values of the parameters for the reference directional coupler are as follows: the length, width and height of both the bus and receiving waveguides are 36.6 μm, 500 nm and 220 nm, respectively; the gap size is 110 nm and the corresponding coupling wavelength of the light is 1550 nm. Both the top and the bottom claddings are the air.

7.2.1. Dimensions of the Cross-section of the Waveguides

For the free-standing strip waveguides of the directional coupler, the two geometrical parameters are the width and the height of the cross-section of the waveguides, which have direct influences on the effective refractive indices. The errors of the dimensions originate from the fabrication processes. For example, the width may not be exactly 500 nm due to the proximity effect. Besides, it has been provided by the manufacturer of the SOI wafers used for the fabrication that the tolerance of thickness of the device layer

is ±10 nm. In the following, the errors will be added to the width and the height of the waveguides respectively.

I. Error of the Width: Fig. 7.9 plots the influence of the errors of the width on the effective refractive indices of the symmetrical and asymmetric modes (n_s and n_a) and the coupling efficiency κ^2. It should be emphasized that the gap between the waveguides is kept as constant while varying the width of the waveguides. It can be learned from Fig. 7.9 that both the refractive indices decrease with smaller waveguide width, meanwhile the difference between n_s and n_a enlarges resulting in smaller L_c. The shaded area highlights the error range corresponding to 1 dB loss. Evidently, the coupling efficiency is more sensitive to smaller width due to a local minimum at the width error of -70 nm. This is because the coupling length at this point is 17.9 µm, hence the light travels almost a complete cycle that almost all the power returns back to the bus waveguide.

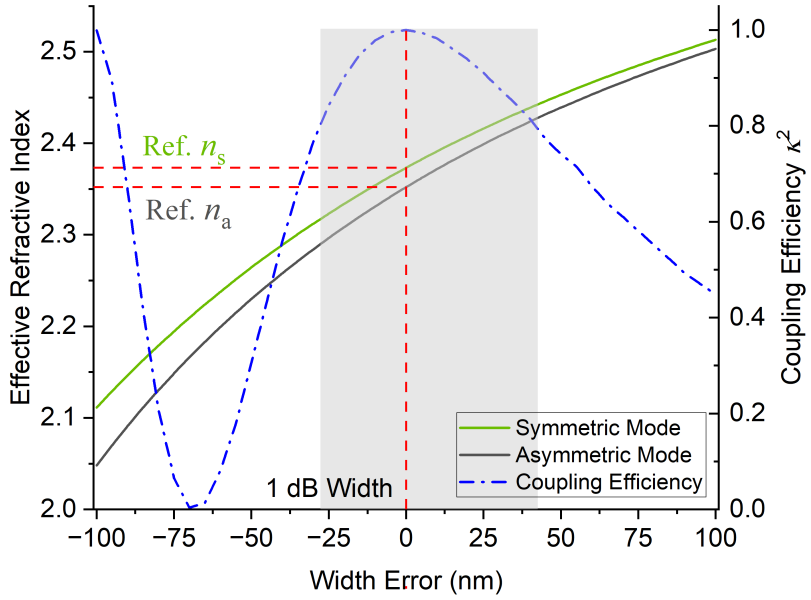


Figure 7.9.: Relations of the symmetric and asymmetric refractive indices n_s and n_a, and the coupling efficiency κ^2 versus the errors of the waveguide width. The shaded area represents the 1 dB width of the error range.

II. Error of the Height: The influence of the height of the waveguide is displayed in Fig. 7.10. In the error range from -20 nm to 20 nm, both n_s and n_a increase nearly linearly with the growing errors in parallel. Therefore, there is minor variation of the coupling length, and the coupling efficiency is insensitive to the errors of waveguide height.

7.2.2. Gap Size

Because the studied directional coupler is a movable structure, this parameter is studied in a wide range from 60 nm to 600 nm, as illustrated in Fig. 7.11. With the increasing gap

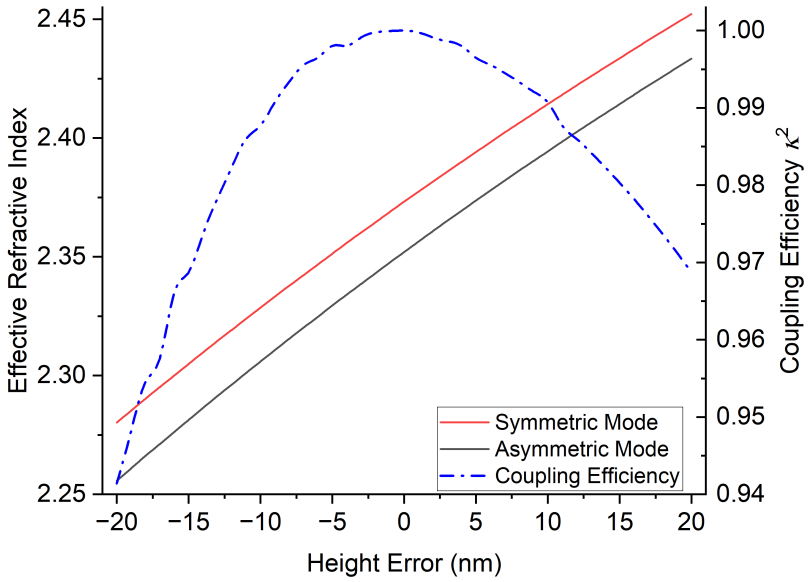


Figure 7.10.: Relations of the symmetric and asymmetric refractive indices n_s and n_a, and the coupling efficiency κ^2 versus the errors of the waveguide height.

size, the symmetric effective refractive index drops significantly in comparison with the asymmetric term, which gradually increases and approaches the symmetric counterpart. The coupling efficiency shows high sensitivity to the gap size, where the upper limit of the 3 dB loss range is 178 nm, as marked out by the shaded area. This suggests that, starting from complete coupling, a displacement of only 100 nm can lead to significant attenuation of the output signal from the receiving waveguide.

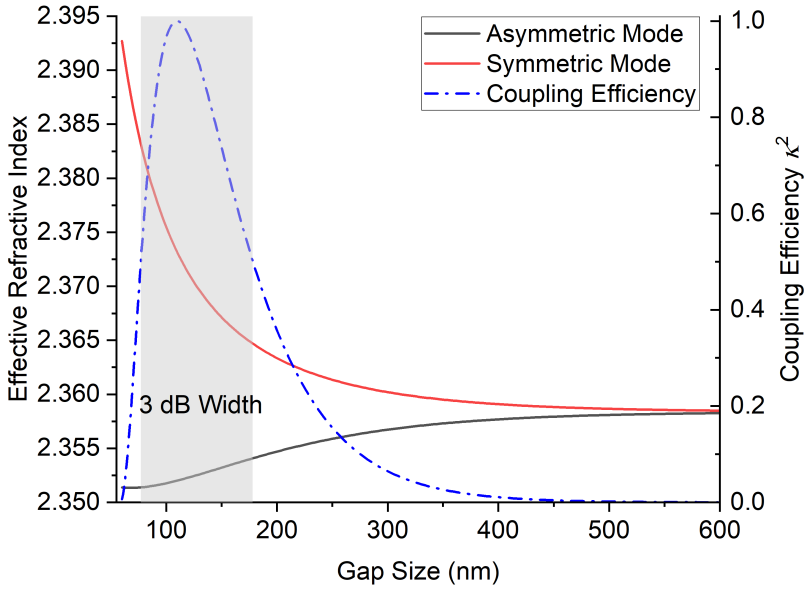


Figure 7.11.: Relations of the symmetric and asymmetric refractive indices n_s and n_a, and the coupling efficiency κ^2 versus the gap size of the directional coupler. The shaded area represents the 3 dB width of the gap size.

7.2.3. Wavelength

Due to the fabrication errors of the grating couplers, the optimal coupling wavelength may deviate from the designed wavelength of 1550 nm. Therefore, it is necessary to know the coupling efficiency change in a wavelength sweep. In Fig. 7.12, both effective refractive indices decline almost linearly with the increasing wavelength approximately in parallel. Hence, the coupling efficiency witnesses a low diminution of about 1 dB in a large wavelength range between 1460 nm and 1550 nm.

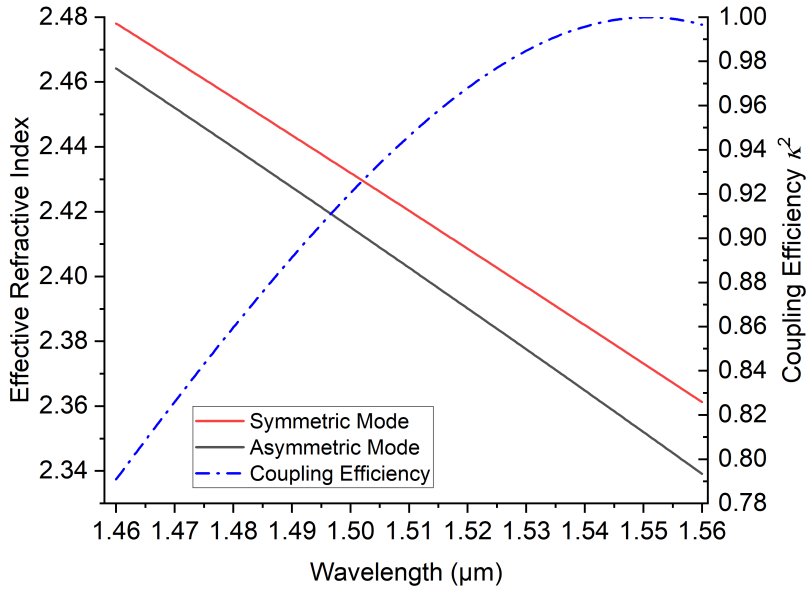


Figure 7.12.: Relations of the symmetric and asymmetric refractive indices n_s and n_a, and the coupling efficiency κ^2 versus the wavelength of the light propagating in the directional coupler.

7.2.4. Refractive Index of Si

The refractive index of Si exhibits a dependency on various factors, including the wavelength [167], the doping level [168–170] and the temperature [171, 172]. For the investigated structure, because of the thermally induced phase transformation of SMA in the nanoactuator, there is a temperature gradient along the free-standing waveguides as well, thus resulting in an increase of the refractive index of Si. Fig. 7.13 presents the influence of the error of the refractive index of Si, Δn_{Si}. The linearly growing effective refractive indices of the two modes gradually approach each other, which gives larger coupling lengths. Consequently, the coupling efficiency decreases, and the loss exceeds 1 dB when the deviation of n_{Si} is larger than 0.144. Based on the thermo-optic coefficient of Si $\mathrm{d}n/\mathrm{d}T = 1.818 \times 10^{-4}\,\mathrm{K}^{-1}$ at the wavelength of 1550 nm [171], this error corresponds to a temperature increase about 790 K, which is significantly higher than required to induce the phase transformation. The influences of the wavelength and the doping level on the refractive index of Si are also low [168, 170, 171], taking in account that the doping level of the device layer of the SOI wafer used for fabrication is about $10^{15}\,\mathrm{cm}^{-3}$. Thereby, it

can be concluded that the factors leading to fluctuations of the refractive index of Si have limited influence on the coupling efficiency.

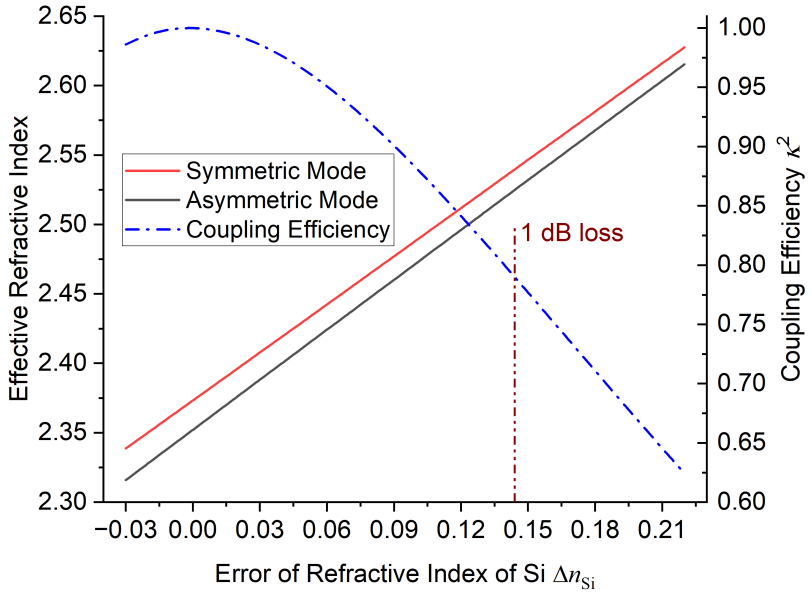


Figure 7.13.: Relations of the symmetric and asymmetric refractive indices n_s and n_a, and the coupling efficiency κ^2 versus the error of the refractive index of Si Δn_{Si}. In the studied error range, the loss is above 1 dB when $\Delta n_{Si} > 0.144$.

7.3. Fabrication Process

The fabrication process of the actuated directional coupler comprises several runs of exposures and etching steps. For clarity of explanation, the major steps are introduced separately in the following.

I. **Chip Preparation**
This fabrication process starts with a SOI chip (Fig. 7.14 (a)). The thickness of the device layer and box layer of these chips are of 220 nm and 3 µm, respectively. Standard cleaning procedures are carried out beforehand and gold markers are patterned on the chips by lift-off prior to structure fabrication due to the demand of more than one exposure step, as introduced in Section 2.2.1.

II. **Definition of Photonic and Mechanical Structures**
The first exposure step is to define the patterns of waveguides, grating couplers, nanoactuators and conductive lines. The negative resist HSQ is coated and patterned by EBL. Different dose values are chosen depending on the feature sizes. These patterns are then transferred to the device layer by RIE with shallowly etched depth of 70 nm (Fig. 7.14 (b)). Etching gases for this process include CHF_3 and SF_6, and the chamber temperature during etching is set at 20 °C. It is important to note that HSQ

is not removed after this dry etching step as it will still act as a hard mask for the following etching step.

III. **Definition of Release Windows**
The second exposure step is to define the release windows, inside which the structures, including nanoactuators and directional couplers, will be free-standing. The positive resist PMMA is coated on the chips and exposed by EBL, as depicted in Fig. 7.14 (c). The pattern transfer is achieved by another cryo-RIE step to fully etch the remaining device layer of 150 nm in the defined windows. Etching gases for this step are SF_6, O_2 and Ar, respectively. The chamber temperature is cooled down to −100 °C. It should be highlighted that there are defined waveguide and actuator patterns in the windows and they are protected by the HSQ from previous exposure step. PMMA is removed after RIE to get prepared for the next wet-etching step (Fig. 7.14 (d)).

IV. **Structure Release**
The chips are immersed in HF-based solution to etch SiO_2 underneath. The etching depth needs to be at least 1 μm to avoid potential interference on waveguides from the substrate. Critical point drying follows afterwards to avoid damage on the structures caused by liquid stiction (Fig. 7.14 (e)). Additionally, HSQ is also dissolved and removed in this wet-etching step due to its composition properties.

V. **SMA Deposition**
Because the SMA layer is required to be locally sputtered on conductive lines and nanoactuators, it is added to the chips by a lift-off process. A bilayer PMMA is coated and exposed to define the sputtering window, then the chips are handed to DC magnetron sputtering of SMA. Similar to the thickness requirement of the bilayer PMMA that having been explained in Section 6.4, it is necessary that the resist is thick enough to fill the height gap between the free-standing structures and the substrate. After lift-off and being dried by critical point drying, the chips are annealed to crystallize SMA (Fig. 7.14 (f)).

The main advantage of this fabrication process is that no additional mask is required during the wet-etching step. Because the waveguides are only partly released, as a comparison, the conventional process is to apply a resist as the protective mask on the waveguides sitting on the substrate. Since HF is highly aggressive to polymer-based resists, the etching time is largely limited and there is risk of introducing contamination from resist. A trial with resist AZ 1505 is taken before the above process is conceived and it is found that the resist starts to peel off during etching. By exploiting the high selectivity between Si and SiO_2 in HF-based solution, shallowly etched rib waveguides function as a perfect mask to define the etching windows.

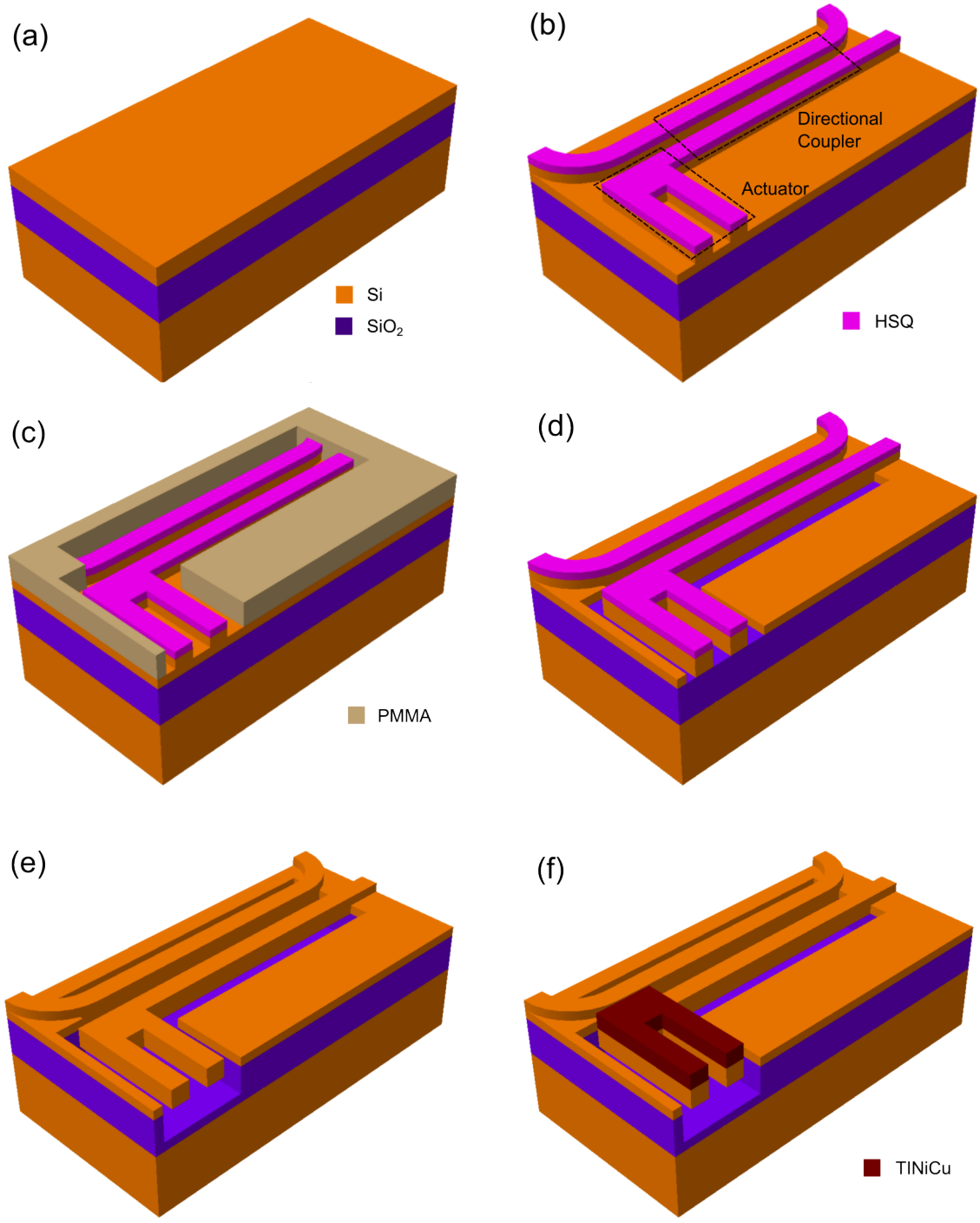
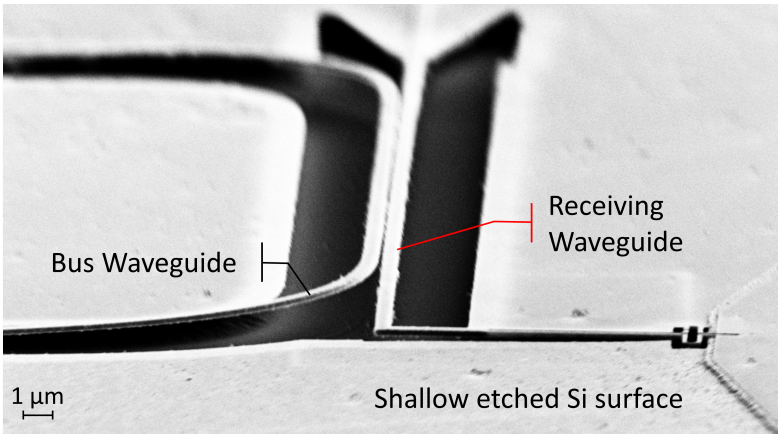


Figure 7.14.: Fabrication processes of SMA-based nanoactuator integrated with the directional coupler. (a) The SOI wafer used for fabrication; (b) The shallow Si RIE of 70 nm. The patterns for this etching step are defined by the first EBL exposure with the resist HSQ. (c) A window for the deep Si RIE is defined by the second EBL exposure with the resist PMMA. (d) Deep Si RIE to expose the box layer in the regions that the structures will be released. (e) Wet-etching of SiO_2 to release the free-standing nanoactuators and directional couplers. (f) 300 nm-thick $Ti_{54}Ni_{31}Cu_{15}$ layer is added to the top of the nanoactuators and the electric circuit (not displayed in the figure) by lift-off. The whole wafer is annealed lastly.

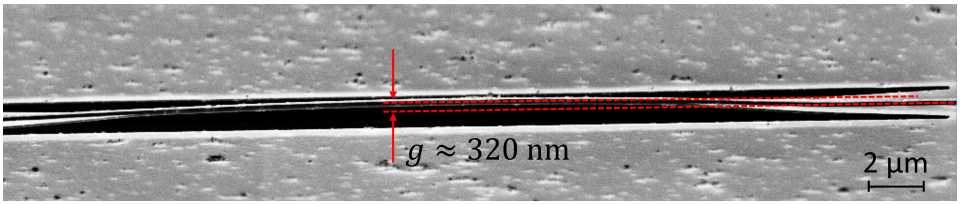
7.4. Fabrication and Experiment Results

7.4.1. Fabrication Results

Firstly, the fabrication results following the process in Section 7.3 are presented. Fig. 7.15(a) shows the free-standing directional coupler integrated with a cantilever beam of 12.5 µm in length after Step (e). The height gap from the shallow etching (Step (b)) and the deep etching (Step (d)) of Si can be clearly seen around the conductive pads at the fixed-end of the cantilever beam. The free-standing structures are released successfully without adhesion problem. However, one can notice a slight height difference between the bus waveguide and the receiving waveguide due to the residual stress in the device layer, which has been discussed in Chapter 6. This is better revealed on the partially released reference waveguide in Fig. 7.15(b). The reference waveguide refers to a waveguide without functional component in order to acquire the transmission loss through the waveguides and the grating couplers. In Fig. 7.15(b), the partially released part of the waveguide is in a shape of the bridge beam of 20 µm in length, and it buckles up with the maximum deflection about 320 nm.



(a) Free-standing directional coupler with a cantilever beam.



(b) Partially released reference waveguide.

Figure 7.15.: SEM images showing (a) a free-standing directional coupler with a 12.5 µm-long cantilever beam and (b) a reference waveguide with a partially released section of 20 µm in length. Both pictures are taken after step (e) of the fabrication process.

Fig. 7.16 presents two identical structures (A and B) at different positions on the same chip after the annealing of $Ti_{54}Ni_{31}Cu_{15}$ layer (Step (f)) in both the side and top views.

The dimensions of both structures are the same as the structure in Fig. 7.15. Due to the severe sidewall deposition during the sputtering process, one can notice that there are the "skirt-like" residuals along the edges of the electric conductive pads. Besides, the isolation gaps between the pads are almost or completely filled causing short-circuit of the current through the double-beam nanoactuators. Another noticeable phenomenon is that the identical nanoactuators deform into different shapes—the beam bends upwards in Structure A while deflects downwards in Structure B. This indicates that the stress distribution is not uniform across the wafer surface. Fig. 7.17 further demonstrates a deformed 10 μm-long nanoactuator, which has a large deflection of about 3.3 μm. In all three examples, the gaps between the directional couplers are significantly enlarged than the designed 110 nm, which causes the light coupling impossible. The unstable and uneven stress distribution across the wafer surface brings considerable challenge of precisely controlling the gap of the directional couplers, which is crucial to the functionality of the actuated directional coupler.

Therefore, in the following only the results of the photonic transmission tests are presented, which are conducted with the photonic circuit after step (e) of the fabrication process.

7.4.2. Setup of Transmission Test

For the fabricated nanophotonic devices, the standard procedure to learn the device performance is to perform transmission tests. A laser beam with a specific power is guided into the integrated photonic circuit from an input optical fiber, which then propagates along the designed waveguide path until it reaches the output terminal connecting to the output optical fiber. The beam from the output end is analyzed to learn its power and phase information, which reveals the loss and performance of on-chip structures. In terms of the directional coupler, the output power from the bus and receiving waveguides are of great interest to learn the coupling efficiency.

Fig. 7.18 shows a sketch of a transmission test platform of integrated photonic chips, which comprises the following different parts.

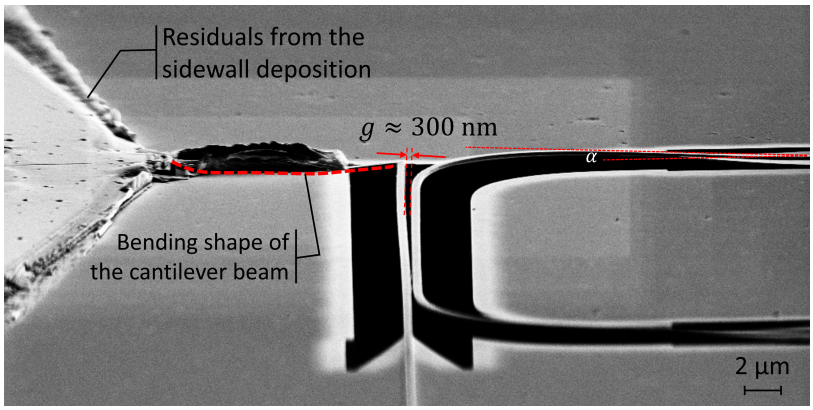
Fiber-coupled laser source: The tunable wavelength of the laser source renders a wavelength sweep feasible, which is necessary to find the optimal functioning wavelength.

Fiber polarization controller: The polarization states of light is converted between the linear, circular and elliptical states by using stress-induced birefringence.

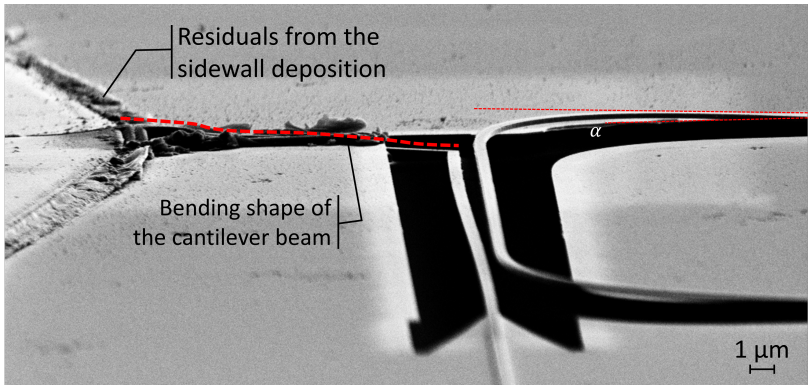
Charge-Coupled Device (CCD) camera: The accurate positions of the fiber tips are monitored in real-time with the help of the magnification lens on the camera.

Optical spectrum analyzers (OSA): The output light from the chip is collected and analyzed.

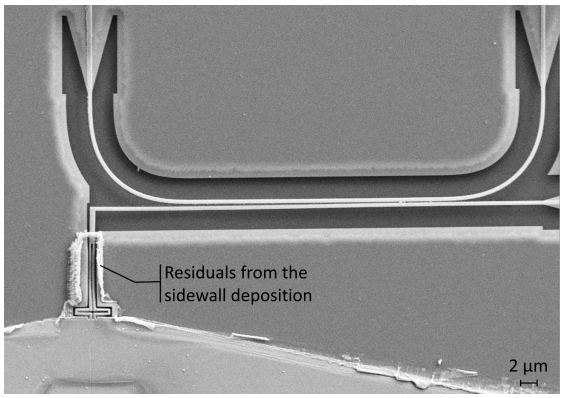
Unless explicitly indicated, the input power from the laser source is set to be 0 dBm for all the transmission tests in the following sections by default.



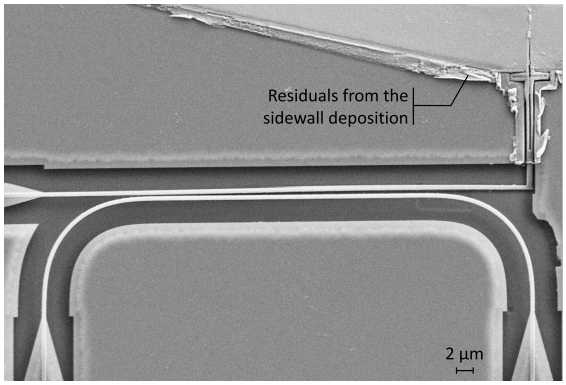
(a) Side view of Structure A.



(b) Side view of Structure B.



(c) Top view of Structure A.



(d) Top view of Structure B.

Figure 7.16.: SEM images of the side and top views of two identical structures with different deformation shapes, which have the same dimensions as the structure in Fig. 7.15.

7.4.3. Transmission Test Results

Reference Waveguides Obtaining the transmission spectrum of a reference waveguide is usually the first step in a transmission test, which gives the information such as the efficiency of the grating coupler, the envelop curve and the loss of a photonic circuit in a

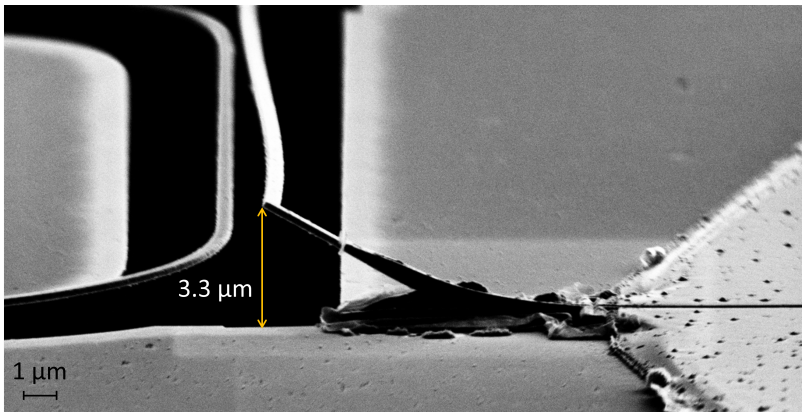


Figure 7.17.: SEM image of a 10 µm-long nanoactuator with a deflection of about 3.3 µm after the annealing of $Ti_{54}Ni_{31}Cu_{15}$ layer.

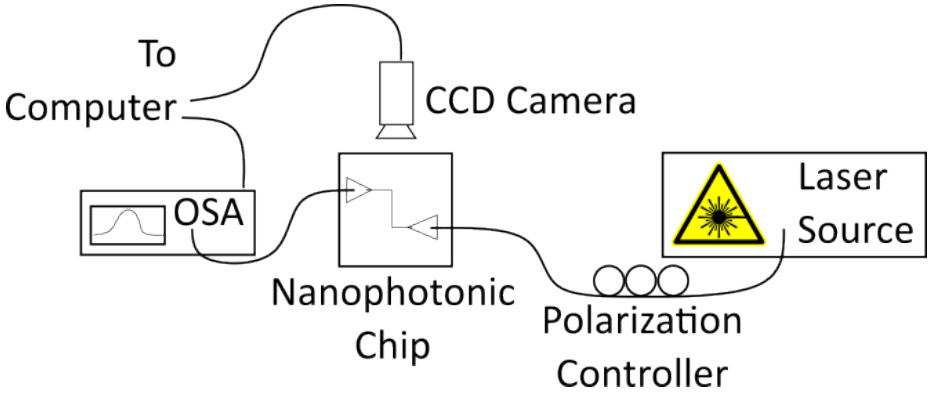


Figure 7.18.: Sketch of the test platform for transmission tests.

certain bandwidth. Fig. 7.19 compares the transmission spectra of two types of reference waveguides: the first one is the pure rib waveguide, and the second one is composed of both rib and strip waveguides by partially releasing a section of the waveguide in the middle. The purpose of comparison is to learn the loss brought by the taper structure. The wavelength range of the tests is between 1460 nm and 1560 nm.

Clearly, the two approximately coincident curves confirm that the loss from the taper structure is negligible. The peaks of the curves fall at the wavelength of 1555 nm, which validates the efficiency of the grating coupler at the incident angle of 10°. The grating coupler also possesses a wide 1 dB bandwidth of about 18 nm (from 1542 nm to 1560 nm). Moreover, the reference waveguides are tested after experiencing the annealing environment of the SMA. The dashed curve in Fig. 7.19 is from a reference rib waveguide before annealing, which indicates that the annealing process of $Ti_{54}Ni_{31}Cu_{15}$ films has minimal influence on the performance of Si photonic circuits.

Free-standing Directional Couplers In Section 7.1, the gap and the length of the directional coupler are designed to be 110 nm and 36.6 µm, respectively, which correspond to the length of full coupling L_c. Fig. 7.20 plots the transmission spectra of the directional couplers

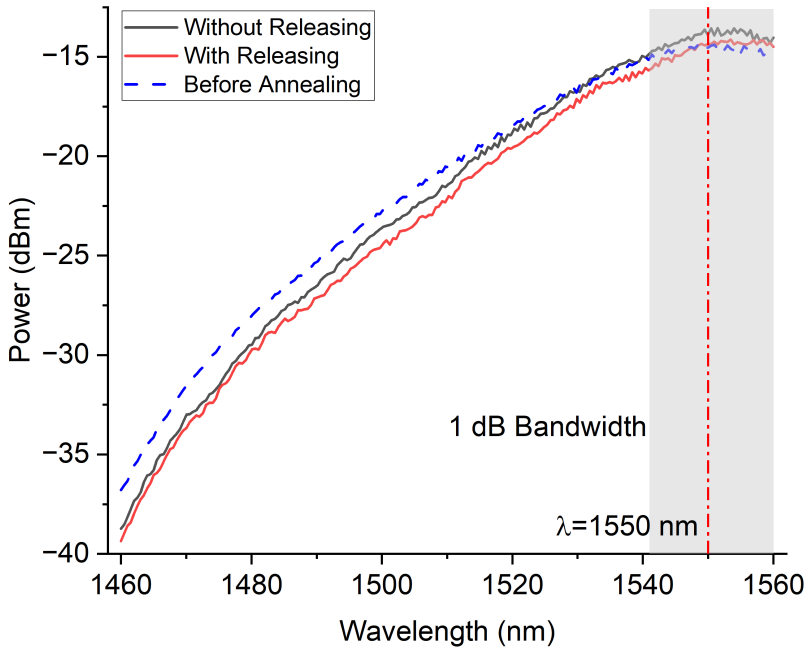
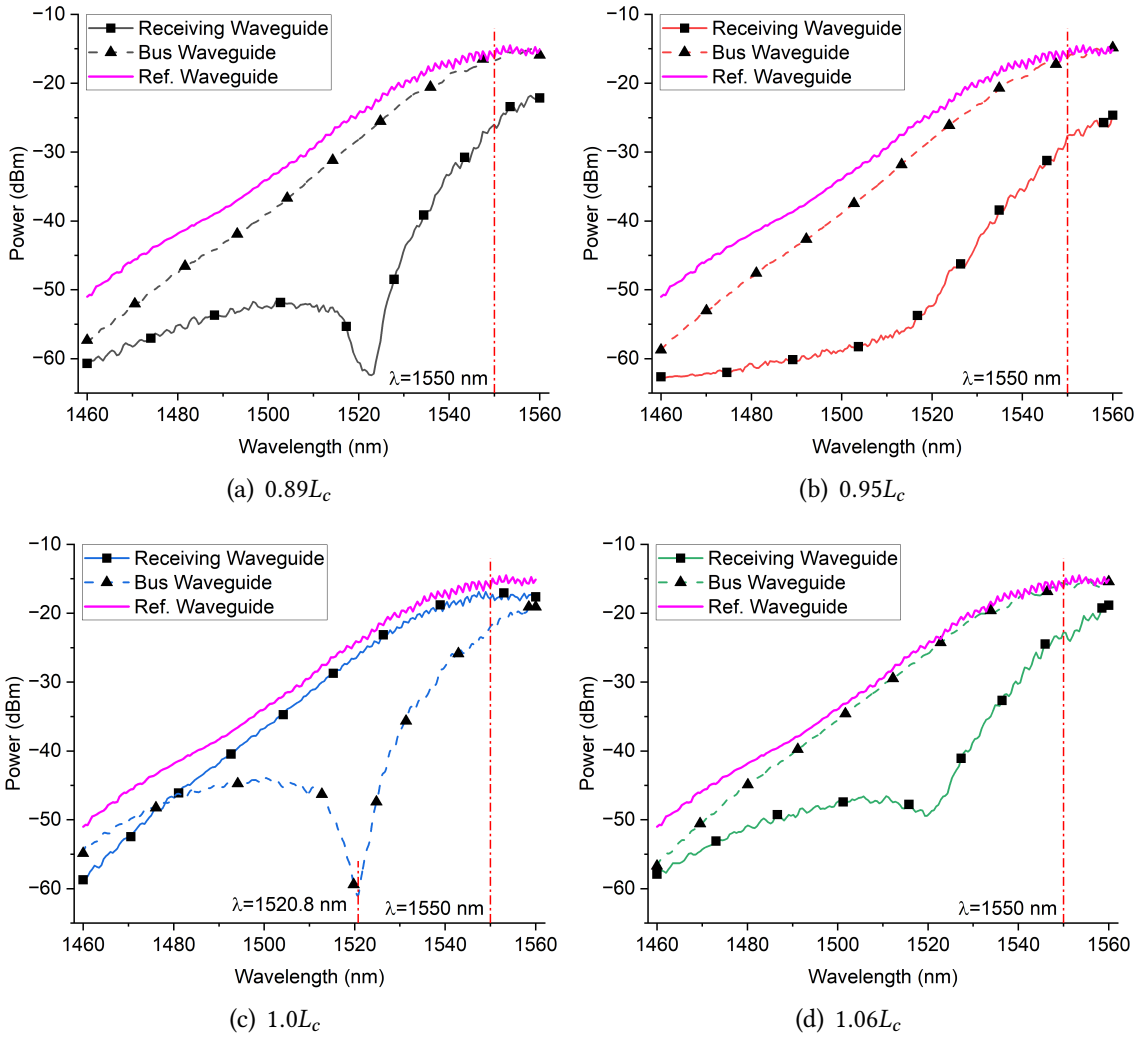


Figure 7.19.: Transmission spectra (solid lines) of the reference waveguides with and without partially releasing section after being annealed. The dashed line is from the reference rib waveguide before annealing. The shaded area represents the 1 dB bandwidth of the grating coupler.

with different lengths ranging from $0.89L_c$ to $1.06L_c$. Evidently, at the wavelength of 1550 nm, it is only at the length of $1.0L_c$ that the output power from the receiving waveguide is higher than the power from the bus waveguide. The exceeded magnitude is about 4 dB corresponding to about 71% of the total power coupled into the receiving waveguide. A full coupling takes place at the wavelength of 1520.8 nm, while it is inconsistent with the analyses in Fig. 7.12. For the other lengths, the power from the bus waveguide is about 9 dB higher on average than the receiving waveguide at the wavelength of 1550 nm. Fig. 7.21 demonstrates the power ratios distributed in the bus and receiving waveguides of different lengths.

The imperfect coupling at the length $1.0L_c$ and the high sensitivity of the coupling efficiency to the length can be attributed to various reasons. One concern is the extra coupling from the non-parallel waveguides (bends) in the neighboring area of the directional coupler. The extra coupling distance can reach to 2.3 μm for a directional coupler designed to be 15.5 μm in length [62]. Due to the systematic error in the exposures, the dimensional deviations of the grating couplers may excite TM mode instead when the phase matching condition can be satisfied. The mode analysis gives that L_c equals to 2.8 μm at TM mode, which is significantly smaller than the length at TE mode. Because of the fragility of the free-standing structure, the variations of the gap size can also contribute to the discrepancy.

Directional Couplers on Box Layer To further verify the analytical and numeric models in the analyses above and exclude the influence of the movable waveguides, the fixed



(a) $0.89L_c$ (b) $0.95L_c$

(c) $1.0L_c$ (d) $1.06L_c$

Figure 7.20.: Transmission spectra of the free-standing directional couplers with various lengths from $0.89L_c$ to $1.06L_c$. The magenta line is the spectrum of the reference waveguide.

directional couplers on box layer with the grating couplers for TM-mode are also fabricated to perform the transmission tests. The dimensions of the waveguides and the gap size are kept the same. A thick PMMA layer is coated on top of the chips as both the protective layer and the top cladding layer. At the wavelength of 1550 nm, the symmetric and asymmetric TM-modes of such directional couplers are presented in Fig. 7.22, in which the color contour represents the value of the magnetic field component in $x-$direction $\vec{H}_x$. The corresponding coupling length L_c is 4.5 µm. As examples, Fig. 7.23 plots the transmission spectra of the directional couplers with lengths of 19.6 µm and 31.3 µm, respectively. The former gives the full coupling at the wavelength of 1556.5 nm while for the latter, the wavelength of full coupling shifts to 1540.5 nm.

At the wavelength of 1550 nm, the power outputs from the receiving and the bus waveguides of a series of lengths are marked out in Fig. 7.24. Taking the extra coupling length

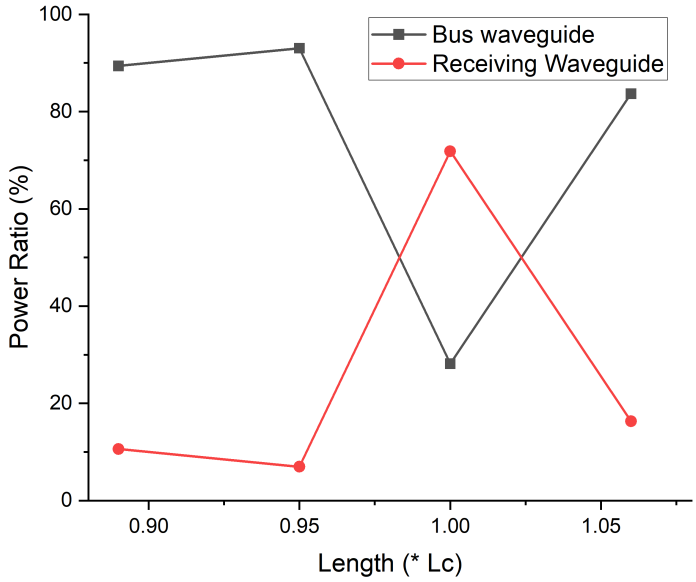


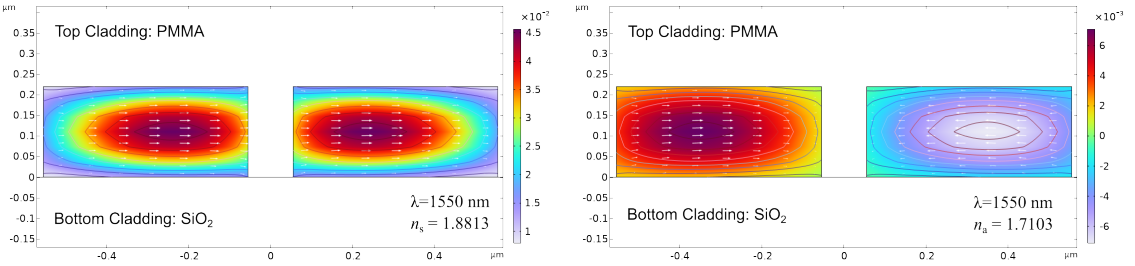
Figure 7.21.: Power ratios in the bus and receiving waveguides of the lengths in Fig. 7.20.

z_{extra} from the non-parallel waveguides into account, the coupling efficiency of Eq. 3.27 can be rewritten as

$$\kappa^2 = \sin^2\left[\frac{\pi}{2}\frac{(z + z_{\text{extra}})}{L_c}\right]. \tag{7.5}$$

Accordingly, the power ratio through the bus waveguide is

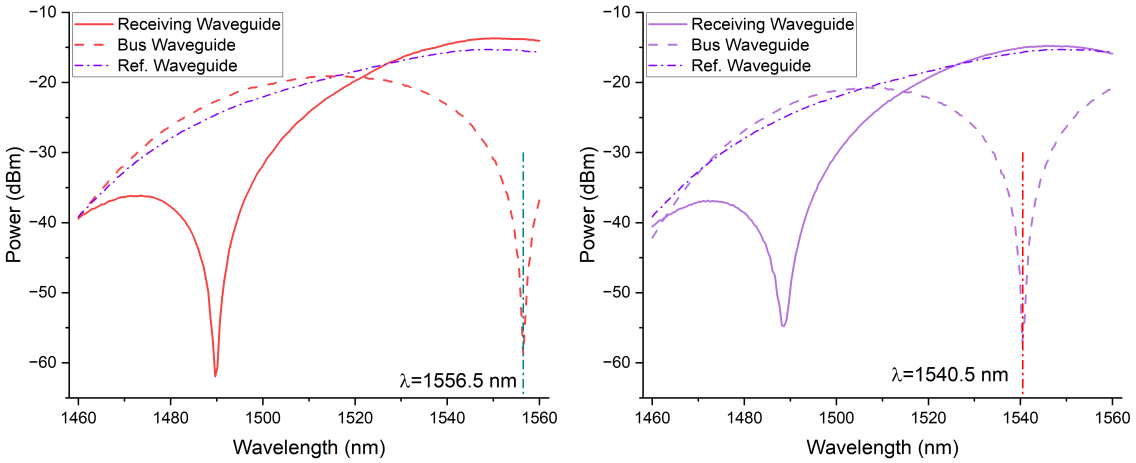
$$t^2 = 1 - \kappa^2 = \cos^2\left[\frac{\pi}{2}\frac{(z + z_{\text{extra}})}{L_c}\right]. \tag{7.6}$$



(a) Symmetric Mode (b) Asymmetric Mode

Figure 7.22.: (a) Symmetric and (b) asymmetric TM-modes of the directional coupler composed by two strip waveguides with a gap of 110 nm at the wavelength of 1550 nm. The top and bottom claddings are PMMA and SiO_2, respectively.

By fitting the data points in Fig. 7.24 with Eqs. 7.5–7.6, both fitted equations give the coupling length of 5.1 µm, which is comparable to the theoretical value 4.5 µm. The fitted extra coupling length $z_{\text{extra}} = 5.25$ µm, which indicates that the coupling effects from non-parallel waveguide bends are not negligible when designing the directional couplers.



(a) 19.6 µm (b) 31.3 µm

Figure 7.23.: Transmission spectra of the fixed directional couplers in TM mode with lengths of (a) 19.6 µm and (b) 31.3 µm.

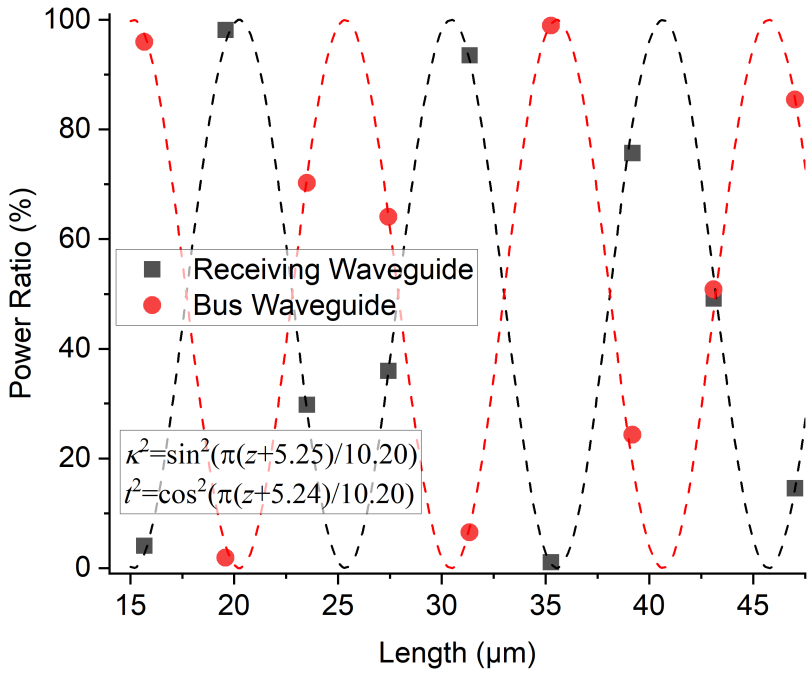


Figure 7.24.: The ratios of power outputs from the receiving and the bus waveguides of different lengths at the wavelength of 1550 nm. The dots are experimental data, and the dashed lines are curves of the fitted functions.

7.5. Discussion

The fabrication results and the transmission spectra of the photonic devices have been presented, which in principle verify the feasibility of the integration of the SMA-based nanoactuator and the directional coupler. Meanwhile, the quality problems and the discrep-

ancy exposed in the fabrication and experiments form the basis for the future amelioration, which will be discussed in detail hereinafter.

I. Stress Engineering in Layers: The stresses in the SMA and Si layers play a crucial role in the deformation behavior of the nanoactuator. It has been revealed that the inhomogeneous stress distribution across the wafer surface seriously impaired the functionality of the actuated directional coupler. Apart from the residual stress in the SMA layer introduced during sputtering and annealing processes, particular attention should be paid to the internal stress in Si layer generated during the fabrication process. Fig. 7.25 shows a directional coupler without nanoactuator. After releasing, both waveguides bend upwards with an exceptionally large deflection of about 5.6 μm, which is not seen in Fig. 7.15(a), even though the same releasing processes are followed. Therefore, to optimize the performance of this structure, it is imperative to pinpoint stress sources and implement precise control of it. In addition, by referring to Eq. 4.17, a stress compensation layer with appropriate residual stress can be added on top of the SMA layer so that the nanoactuator can start from a flat state at room temperature.

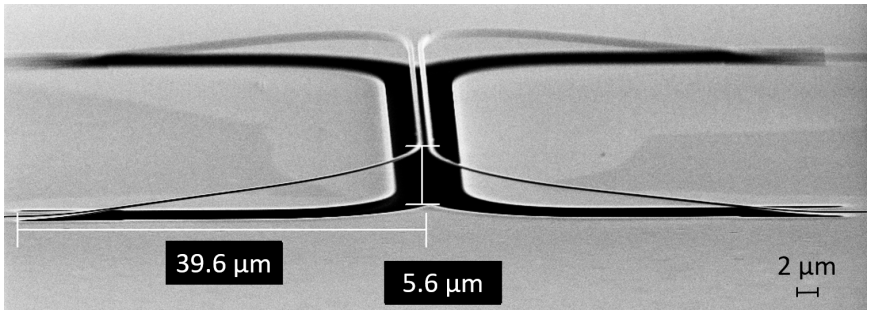


Figure 7.25.: A released directional coupler without the nanoactuator. Though there is only the Si layer, the directional coupler has a deflection of about 5.6 μm.

II. Directionality of the SMA Sputtering Process: Fig. 7.16 explicitly demonstrates the influences of the sidewall deposition. Besides the short-circuit problem, the sidewall depositions not only change the original designed dimensions, but also bring harmful metal particles during the lift-off process. The remaining particles on the photonic circuit can greatly attenuate the light propagation. The adjustments on the sputtering parameters and processing with IBE after sputtering and before lift-off are necessary to mitigate the negative influences of sidewall deposition.

III. Structural Robustness: Another reason contributing to the difficulty of deflection controlling is the large length of the directional coupler. Although the TE-mode has higher effective refractive indices to "trap" the light, the required coupling length of 36.6 μm also causes the mechanical sensitivity of the structure to external disturbance. Integrating tapers into the design [173] or switching to TM-mode are two potential methods to improve the structural robustness with smaller coupling length. Meanwhile, the degradation of the

deflection change of the nanoactuator, which is defined by Eq. 7.4, and the sensitivity of the coupling efficiency with shorter length require an overall consideration to guarantee both the mechanical and optical performance of this nanoactuator integrated directional coupler.

7.6. Summary

In this chapter, the design, fabrication and photonic characterization of the nanoactuator integrated directional couplers are investigated analytically and experimentally. The multidisplinary structure is decomposed into the photonic circuits, the mechanical actuators and the electric circuits for separate designs. The key components, directional couplers, functions in TE-mode with the coupling length of 36.6 μm at the wavelength of 1550 nm. The parametric studies of the directional coupler reveal that the width of the waveguides and the gap size have more substantial influences on the coupling efficiency compared to other parameters. In the fabrication process that requires three exposure steps, the combination of shallow etching and deep etching of Si renders more controllable wet etching process of SiO_2. The fabrication results show that stress engineering of both SMA and Si layers plays an important role in precise controlling of the gap size. About 71% of the input power is successfully coupled into the receiving waveguide at the designed length by performing the transmission tests. The tests with fixed directional couplers verify that the deviations between the analytical and experimental results of the free-standing structures can be attributed to the error of the gap size control.

8. Summary and Outlook

8.1. Summary

Belonging to the category of smart materials, shape memory alloys (SMAs) exhibit high work density and potential scalability of dimensions down to nanometer size, which renders SMAs highly attractive for applications in the fields of MEMS and NEMS. Integrating SMA films into the design of bistable structures and MEMS-tuned photonic circuits brings new possibilities of having more compact and energy efficient nano-scale actuators for broader application domains, e.g., electric relay, photonic switch and non-volatile memory.

The overall objective of this dissertation is to perform bistable out-of-plane actuation at the nano-scale enabling novel tunable devices for optic and photonic applications. To achieve this objective, SMA-based bilayer and multilayer cantilevers are investigated including their design, fabrication and characterization. The base SMA materials are $Ti_{54}Ni_{31}Cu_{15}$ films developed by the collaboration partners at RUB. The first part of the dissertation addresses the development of analytical models to predict the mechanical and thermal performances of multilayer beam structures. This allows the model-based development of Ti-Ni-Cu/Si bilayer cantilevers and advanced bistable nanoactuators with four-beam structure. Finally, the potential use of the nanoactuators in mechanically tunable photonic directional couplers is explored. Accordingly, appropriate fabrication processes are developed for the bistable nanoactuators and the mechanically tuned photonic directional couplers. Various requirements of techniques are taken into account, including EBL, RIE and magnetron sputtering. The major results from this work are summarized in the following.

I. Analytical models predicting the mechanical and thermal performances of multilayer beam structures

Based on the Euler-Bernoulli beam theory, an analytical mechanical model is developed to calculate the stress distribution, curvature and deflection of multilayer beam structures deformed due to the inelastic strains in the layers, such as thermal and residual strain. This concise model does not rely on the assumption of small deflection and is compatible with the non-linear material properties present in SMAs.

Starting from the Fourier's law, an analytical thermal model is derived to describe the temperature profile and heat dissipation mechanism of Joule-heated multilayer beam structures, which can be in cantilever or bridge shape. It is found that at the scale below

micrometer, heat conduction to the substrate plays the major role in the heat dissipation mechanism and geometrical dimensions have negligible influence on the maximum temperature on the beam. This model is validated experimentally by a Joule-heated cantilever beam at millimeter scale.

Despite the limitations on the applicability of material and structural shape, these two analytical models are convenient for having a fast overall understanding of the involved parameters and preliminary designs of mulitlayer nanoactuators. The models also form the bases of the further investigations of this work.

II. Mechanical and thermal analyses of Ti-Ni-Cu/Si bilayer nanoactuators

The phase transformation temperatures of 600 nm magnetron sputtered $Ti_{54}Ni_{31}Cu_{15}$ film on Si substrate are determined by temperature-dependent electrical resistance measurement, and the result shows that the film transforms above room temperature with a narrow hysteresis width of 1.9 °C. This confirms the feasibility of the $Ti_{54}Ni_{31}Cu_{15}$ films for the applications in actuation. On this basis, the electro-thermo-mechanical performances of the Joule-heated SMA/Si bilayer cantilever nanoactuators are investigated analytically and numerically. The mechanical model reveals that the deflection evolution with hysteresis in a thermal cycle can be attributed to several competitive effects, i.e., the mismatch of thermal expansion coefficients effect, the ΔE-effect and shape memory effect. The stroke of the nanoactuator reaches the maximum when the thickness of $Ti_{54}Ni_{31}Cu_{15}$ layer is half of the Si layer. Due to the temperature gradient disclosed by the thermal model, an appropriate current range is proposed that 80% of beam length can complete phase transformation to austenite state while overheating is avoided.

Folding the beam length that partially transforms into a wing structure is an effective approach to mitigate the temperature gradient. The diminution of mechanical stiffness brought by the wing structure is studied by employing the Castigliano's second theorem. Combining with the joint electro-thermo-mechanical simulation, it is found that the optimal range of the wing ratio is 20%–25% to attain a compromise among the stroke, power consumption and structural stiffness.

III. Bistable out-of-plane SMA nanoactuators with four-beam design

The novel design is composed of four cantilever beams arranged in parallel. The $Ti_{54}Ni_{31}Cu_{15}$ film is selectively deposited onto the two center beams, which leads to buckling of the two outside beams because of the residual stress in $Ti_{54}Ni_{31}Cu_{15}$ film. The thermal effect and shape memory effect provide the moment to switch between the two stable buckling states. The structure of the nanoactuator is equivalent to a cantilever beam supported with finite stiffnesses at the free-end for the buckling analyses. Precise adjustments of the geometrical dimensions are essential to enlarge the difference between the first and second mode of critical buckling forces.

The fabrication process successfully demonstrates the feasibility of selective deposition of SMA layers on free-standing structures. The deflection of the fabricated structures without SMA layer suggests that there is significant position-dependent internal stress in the Si layer, which can be decomposed to a compressive term resulting in buckling and a gradient term resulting in a bending moment. Snapping with a stroke of 1.67% beam length is observed during the nanoindention test of structures without SMA layer. After the deposition and annealing of SMA layer, the initial deflections of the nanoactuators further increase. With specific geometrical dimensions, the nanoactuator snaps during the nanoindention test with an enhanced stroke about 3.6% of beam length, accompanied by a force surge of around 98.6 μN. The non-linear force-displacement relation demonstrates the requirement of additional energy input during the snapping process. The linear loading process before snapping indicates the extra moment from residual stress that the switching moment should overcome before entering the snapping stage. These results mark milestones in the development of bistable nanoactuators based on the buckling induced by internal stress.

IV. Integration of SMA nanoactuators with photonic directional couplers

Based on the Ti-Ni-Cu/Si bilayer nanoactuators that have been developed, the photonic directional couplers are designed to be mechanically tuned by the nanoactuators. Because of the free-standing structures, there is a change of the bottom cladding materials between SiO_2 and air, resulting in propagation loss due to mode mismatch. Appropriate taper structures are designed as transitions between the strip waveguide (air bottom cladding) and the rib waveguide (SiO_2 bottom cladding) to minimize the propagation loss. All photonic components are studied based on the TE-mode. Special attention is paid to the mechanical influence of the waveguide on the nanoactuator. It is analytically proven that the constraining of waveguide has minor influence on the deflection of nanoactuator, if the length of the free-standing waveguide is far larger than the length of the nanoactuator.

The developed fabrication process comprises three lithography steps, i.e., two steps defining patterns of the rib waveguide and the strip waveguide and the last step defining the deposition window of $Ti_{54}Ni_{31}Cu_{15}$ film. The chemical nature of HSQ as a silicon oxide compound is fully exploited in the process, as the high selectivity of HSQ renders it to survive two steps of Si RIE with different recipes. In addition, the rib waveguide naturally functions as a hard mask with high selectivity in the wet etching of SiO_2. Utilizing HSQ resist and adopting the rib waveguide design are proved to significantly simplify the fabrication process and avoid contamination. The process of selectively depositing SMA layer on free-standing structures from the fabrication of bistable nanoactuators is also successfully integrated.

Transmission tests of the reference waveguides confirm that the designed taper structures and the annealing process of $Ti_{54}Ni_{31}Cu_{15}$ film bring negligible additional propagation loss to photonic circuits. Because of the internal stress in the Si layer and the observed severe sidewall deposition of SMA layer, the fabricated nanoactuators are highly degraded. Nevertheless, 71% of light power is successfully coupled into the neighboring waveguide

at the designed length of a free-standing directional coupler. Fixed directional couplers are also fabricated for validating the simulation model. The experimental coupling length of 5.1 μm is comparable to the theoretical length of 4.5 μm. In addition, there is an extra coupling length of 5.25 μm coming from the non-parallel waveguide bends connecting the directional couplers to the grating couplers, which should be highlighted for the design of directional couplers with precise coupling control.

8.2. Outlook

These results open new possibilities of designing bistable out-of-plane SMA nanoactuators and mechanically tuned photonic devices by employing SMA nanoactuators. Although the principal designs, fabrication processes and characterizations have been performed in this dissertation, the precise controlling of stress, deformation mechanism and influential parameters in fabrication still need further investigations and optimizations. In detail, it would be valuable to explore the following aspects in the future.

I. Stress management of Si layer: origin, control and utilization

The observed internal stress in the Si layer significantly deviates the shapes of the fabricated structures from the original design, which is witnessed in both the bistable nanoactuators and the free-standing directional couplers. The primary step to resolve this issue is identifying the generation mechanism of the stress. Having confirmed that this position-dependent stress is generated during the process of SiO_2 wet etching, it is necessary to systematically examine the individual influence of each adjustable variable in this process through the control variable method. The modifiable variables include the type of etchant, etching time, DI-water rinsing time, spin and exchange rates during critical point drying. In addition, implementing techniques such as Raman spectroscopy is another possible approach to determine the stress distribution in the deformed Si free-standing structures.

Having understood the origin of the internal stress in Si layer allows for the deliberate manipulation of the stress, leading to the improved performances of the designed devices. Precisely, the uniform compressive component of the stress is beneficial to creating buckling in the bistable design, whereas the proportional term, which gives a bending moment, causes undesirable initial deformation hence should be suppressed.

II. Improvement of the SMA film quality

The fabricated bilayer structures highlight two aspects that require the improvements of the SMA film quality, i.e., sidewall deposition and stress uniformity. Systematical tuning and optimizing the sputtering parameters, such as chamber pressure and sputtering power, assist in the amelioration of both problems. IBE technique is an effective post-processing tool to mitigate the problem of sidewall deposition, while the damages caused to the

surfaces because of low selectivity need to be taken into account. The annealing condition plays an important role in the formation of the residual stress in the SMA layer, providing further possibilities of tuning the stress distribution.

III. Buckling mechanism and moment tuning in the design of bistable out-of-plane nanoactuators

Although snapping phenomenon is witnessed during the loading process of the bistable SMA nanoactuator, the position presenting the largest stroke locates at the middle of beam, posing difficultly in the application of the nanoactuator. It is preferred that the maximum stroke takes place at the free-end of the cantilever-shape nanoactuator, allowing the potentially connected passive devices, e.g., the photonic directional coupler, to achieve larger motion stroke. It has been demonstrated that the buckling shape undergoes changes with the variations of specific geometrical parameters. More work is required to fully understand the influences of the geometrical dimensions and buckling stress on the buckling shape and snapping stroke by simulation and experiment.

Beyond this, enlarging the gap size between the nanoactuator and the SiO_2 substrate also facilitates larger snapping stroke. Extending the time of SiO_2 wet etching is a straightforward solution, whereas it may further increase the stress level in the Si layer and aggravate the unintended under-etching beneath patterns such as electric conductive pads. A more promising approach is to insert a step of highly anisotropic SiO_2 RIE before the step of wet etching. Two more adjustments of the fabrication process are required accordingly: firstly, another lift-off of the hard mask for defining the patterns of bistable structures is required to replace the currently used negative resist. Secondly, the bilayer PMMA thicknesses and exposure doses for the lift-off of $Ti_{54}Ni_{31}Cu_{15}$ film need to be adjusted as well.

The residual moment induced by internal and residual stress in the Si and SMA layers is a tremendous obstacle to achieving bistability, because this moment elevates the moment threshold that the switching moment needs to cross. It is advantageous to lower the residual moment (ideally to zero) and increase the moment for state switching, i.e., the moment from thermal effect and shape memory effect. Apart from precisely manipulating the stresses in the Si and SMA layers, another feasible method to lower the residual moment is to introduce a third compensation layer above the SMA layer, which should possess similar stress distribution to the Si layer. To avoid an additional lithography and lift-off step, this layer needs to be deposited immediately after the SMA sputtering. Corresponding, the material of the third layer should be compatible with the annealing condition of SMA without degrading it, i.e., considerable inter-diffusion during annealing. Potential candidates of the third layer material include SiO_2 and Si_3N_4.

Increasing the thermal moment can be achieved by tuning the phase transformation of SMA to show relatively high transformation temperatures and narrow hysteresis width. On the one hand, high A_s leads to larger thermal moment in the martensitic state; on the other hand, low A_f reduces the possibility of overheating.

IV. Mechanical robustness of directional couplers

Precision in the gap size control plays the key role in the coupling efficiency of directional couplers tuned by nanoactuators. The initial inaccuracy of gap size observed in this work can be improved by stress management in the Si layer, which has been discussed above. Enhancing the mechanical stiffness of the directional couplers renders the structure to be less prone to external perturbations during functioning, and a favorable approach is to redesign the directional coupler from TE-mode to TM-mode, owing to the considerable decrease of full coupling length from 36.6 μm to 2.8 μm. Accordingly, the grating couplers need to be switched to the TM-mode type. Combination with the designed bistable nanoactuator is also beneficial to promoting structural robustness due to the tendency of staying at stable state of bistable structure.

Bibliography

[1] Pritha Ghosh, Ashwin Rao, and Arun R. Srinivasa. "Design of multi-state and smart-bias components using Shape Memory Alloy and Shape Memory Polymer composites". In: *Materials & Design* 44 (Feb. 2013), pp. 164–171. DOI: 10.1016/j.matdes.2012.05.063.

[2] Ashwin Rao, A. R. Srinivasa, and J. N. Reddy. "Introduction to Shape Memory Alloys". In: *Design of Shape Memory Alloy (SMA) Actuators*. Cham: Springer International Publishing, May 2015. Chap. 1, pp. 1–31. DOI: 10.1007/978-3-319-03188-0_1.

[3] William Smith and Javad Hashemi. *Foundations of materials science and engineering*. Seventh edition. New York, NY: McGraw Hill, 2023.

[4] P.K. Kumar and D.C. Lagoudas. "Introduction to Shape Memory Alloys". In: *Shape Memory Alloys: Modeling and Engineering Applications*. Ed. by D.C. Lagoudas. Boston, MA: Springer US, Mar. 2008. Chap. 1, pp. 1–51. DOI: 10.1007/978-0-387-47685-8_1.

[5] Vincenza Antonucci and Alfonso Martone. "Phenomenology of shape memory alloys". In: *Shape Memory Alloy Engineering: For Aerospace, Structural, and Biomedical Applications*. Ed. by Antonio Concilio et al. Second Edition. Boston: Butterworth-Heinemann, Jan. 2021. Chap. 4, pp. 115–139. DOI: 10.1016/B978-0-12-819264-1.00004-2.

[6] M. Kohl et al. "Ni-Mn-Ga shape memory nanoactuation". In: *Applied Physics Letters* 104.4 (Jan. 2014), p. 043111. DOI: 10.1063/1.4863667.

[7] Aditya Chauhan et al. "A review and analysis of the elasto-caloric effect for solid-state refrigeration devices: Challenges and opportunities". In: *MRS Energy & Sustainability* 2.1 (Dec. 2015), p. 16. DOI: 10.1557/mre.2015.17.

[8] Al Arsh Basheer. "Advances in the smart materials applications in the aerospace industries". In: *Aircraft Engineering and Aerospace Technology* 92.7 (June 2020), pp. 1027–1035. DOI: 10.1108/AEAT-02-2020-0040.

[9] Takashi Mineta and Yoichi Haga. "Materials and Processes in Shape Memory Alloy". In: *MEMS Materials and Processes Handbook*. Ed. by Reza Ghodssi and Pinyen Lin. Boston, MA: Springer US, Jan. 2011. Chap. 6, pp. 355–402. DOI: 10.1007/978-0-387-47318-5_6.

[10] Ashwin Rao, A. R. Srinivasa, and J. N. Reddy. "Manufacturing and Post Treatment of SMA Components". In: *Design of Shape Memory Alloy (SMA) Actuators*. Cham: Springer International Publishing, May 2015. Chap. 3, pp. 43–60. DOI: 10.1007/978-3-319-03188-0_3.

[11] Ferdinando Auricchio et al. "SMA biomedical applications". In: *Shape Memory Alloy Engineering*. Ed. by Antonio Concilio et al. Second Edition. Boston: Butterworth-Heinemann, Jan. 2021. Chap. 19, pp. 627–658. DOI: 10.1016/B978-0-12-819264-1.00019-4.

[12] Hao Yang et al. "A lightweight prosthetic hand with 19-DOF dexterity and human-level functions". In: *Nature Communications* 16.1 (Jan. 2025), p. 955. DOI: 10.1038/s41467-025-56352-5.

[13] S. Miyazaki, Y. Q. Fu, and W. M. Huang. "Overview of sputter-deposited TiNi based thin films". In: *Thin Film Shape Memory Alloys: Fundamentals and Device Applications*. Ed. by Shuichi Miyazaki, Yong Qing Fu, and Wei MinEditors Huang. Cambridge University Press, 2009. Chap. 1, pp. 1–72. DOI: 10.1017/CBO9780511635366.

[14] Yves Bellouard. "Shape memory alloys for microsystems: A review from a material research perspective". In: *Materials Science and Engineering: A* 481-482 (May 2008), pp. 582–589. DOI: 10.1016/j.msea.2007.02.166.

[15] Manfred Kohl et al. "SMA Foils for MEMS: From Material Properties to the Engineering of Microdevices". In: *Shape Memory and Superelasticity* 4.1 (Mar. 2018), pp. 127–142. DOI: 10.1007/s40830-017-0144-x.

[16] Sanaz Rastjoo et al. "Development and co-integration of a SMA/Si bimorph nanoactuator for Si photonic circuits". In: *Microelectronic Engineering* 225 (Mar. 2020), p. 111257. DOI: 10.1016/j.mee.2020.111257.

[17] Jan Frenzel et al. "High-performance elastocaloric materials for the engineering of bulk- and micro-cooling devices". In: *MRS Bulletin* 43.4 (Apr. 2018), pp. 280–284. DOI: 10.1557/mrs.2018.67.

[18] Joel Joseph et al. "Upscaling of Thermomagnetic Generators Based on Heusler Alloy Films". In: *Joule* 4.12 (Dec. 2020), pp. 2718–2732. DOI: 10.1016/j.joule.2020.10.019.

[19] Wangshu Zheng et al. "Design of Shape Memory Ceramics: Principles, Strategies and Perspectives". In: *JOM* 76.7 (May 2024), pp. 3604–3621. DOI: 10.1007/s11837-024-06639-7.

[20] Nicola Manca et al. "Planar Nanoactuators Based on VO_2 Phase Transition". In: *Nano Letters* 20.10 (Oct. 2020), pp. 7251–7256. DOI: 10.1021/acs.nanolett.0c02638.

[21] Run Shi et al. "Recent advances in fabrication strategies, phase transition modulation, and advanced applications of vanadium dioxide". In: *Applied Physics Reviews* 6.1 (Mar. 2019), p. 011312. DOI: 10.1063/1.5087864.

[22] Rafmag Cabrera, Emmanuelle Merced, and Nelson Sepúlveda. "Performance of Electro-Thermally Driven VO_2-Based MEMS Actuators". In: *Journal of Microelectromechanical Systems* 23.1 (Feb. 2014), pp. 243–251. DOI: 10.1109/JMEMS.2013.2271774.

[23] José Figueroa et al. "VO_2-based micro-electro-mechanical tunable optical shutter and modulator". In: *Optics Express* 29.16 (Aug. 2021), pp. 25242–25253. DOI: 10.1364/OE.428165.

[24] Michael E. A. Warwick and Russell Binions. "Advances in thermochromic vanadium dioxide films". In: *Journal of Materials Chemistry A* 2.10 (Mar. 2014), pp. 3275–3292. DOI: 10.1039/c3ta14124a.

[25] Tuan Duc Vu et al. "Physical vapour deposition of vanadium dioxide for thermochromic smart window applications". In: *Journal of Materials Chemistry C* 7.8 (Feb. 2019), pp. 2121–2145. DOI: 10.1039/c8tc05014g.

[26] Ryan M. Briggs, Imogen M. Pryce, and Harry A. Atwater. "Compact silicon photonic waveguide modulator based on the vanadium dioxide metal-insulator phase transition". In: *Optics Express* 18.11 (May 2010), pp. 11192–11201. DOI: 10.1364/OE.18.011192.

[27] Kevin J. Miller et al. "Silicon waveguide optical switch with embedded phase change material". In: *Optics Express* 25.22 (Oct. 2017), pp. 26527–26536. DOI: 10.1364/OE.25.026527.

[28] Zhihua Zhu et al. "Dynamically Reconfigurable Metadevice Employing Nanostructured Phase-Change Materials". In: *Nano Letters* 17.8 (July 2017), pp. 4881–4885. DOI: 10.1021/acs.nanolett.7b01767.

[29] Luis David Sánchez et al. "Experimental demonstration of a tunable transverse electric pass polarizer based on hybrid VO_2/silicon technology". In: *Optics Letters* 43.15 (Aug. 2018), pp. 3650–3653. DOI: 10.1364/OL.43.003650.

[30] Petr Markov et al. "Optically Monitored Electrical Switching in VO_2". In: *ACS Photonics* 2.8 (July 2015), pp. 1175–1182. DOI: 10.1021/acsphotonics.5b00244.

[31] Arash Joushaghani et al. "Wavelength-size hybrid Si-VO_2 waveguide electroabsorption optical switches and photodetectors". In: *Optics Express* 23.3 (Feb. 2015), pp. 3657–3668. DOI: 10.1364/OE.23.003657.

[32] Herman M. K. Wong et al. "Broadband, Integrated, Micron-Scale, All-Optical Si_3N_4/VO_2 Modulators with pJ Switching Energy". In: *ACS Photonics* 6.11 (Oct. 2019), pp. 2734–2740. DOI: 10.1021/acsphotonics.9b00708.

[33] Yunteng Cao et al. "Bistable Structures for Advanced Functional Systems". In: *Advanced Functional Materials* 31.45 (Aug. 2021), p. 2106231. DOI: 10.1002/adfm.202106231.

[34] Bernhard Winzek et al. "Recent developments in shape memory thin film technology". In: *Materials Science and Engineering: A* 378.1 (July 2004). European Symposium on Martensitic Transformation and Shape-Memory, pp. 40–46. DOI: 10.1016/j.msea.2003.09.105.

[35] Roman Vitushinsky, Sam Schmitz, and Alfred Ludwig. "Bistable Thin-Film Shape Memory Actuators for Applications in Tactile Displays". In: *Journal of Microelectromechanical Systems* 18.1 (Feb. 2009), pp. 186–194. DOI: 10.1109/JMEMS.2008.2009816.

[36] Tobias Sterzl et al. "Bistable shape memory thin film actuators". In: *Smart Structures and Materials 2003: Active Materials: Behavior and Mechanics*. Ed. by Dimitris C. Lagoudas. Vol. 5053. SPIE, Aug. 2003, pp. 101–109. DOI: 10.1117/12.484696.

[37] Sabrina M. Curtis et al. "TiNiHf/SiO_2/Si shape memory film composites for bi-directional micro actuation". In: *International Journal of Smart and Nano Materials* 13.2 (May 2022), pp. 293–314. DOI: 10.1080/19475411.2022.2071352.

[38] Yiping Liu et al. "Thermomechanics of shape memory polymers: Uniaxial experiments and constitutive modeling". In: *International Journal of Plasticity* 22.2 (Feb. 2006), pp. 279–313. DOI: https://doi.org/10.1016/j.ijplas.2005.03.004.

[39] Jinsong Leng et al. "Shape-memory polymers and their composites: Stimulus methods and applications". In: *Progress in Materials Science* 56.7 (Sept. 2011), pp. 1077–1135. DOI: https://doi.org/10.1016/j.pmatsci.2011.03.001.

[40] Weiqun Liu et al. "Investigation of a buckled beam generator with elastic clamp boundary". In: *Smart Materials and Structures* 25.11 (Oct. 2016), p. 115045. DOI: 10.1088/0964-1726/25/11/115045.

[41] Dhananjay Yadav et al. "A Two-Terminal Bistable Electrothermally Actuated Microswitch". In: *Journal of Microelectromechanical Systems* 28.3 (June 2019), pp. 540–549. DOI: 10.1109/JMEMS.2019.2904997.

[42] Jin Qiu et al. "A bulk-micromachined bistable relay with U-shaped thermal actuators". In: *Journal of Microelectromechanical Systems* 14.5 (Oct. 2005), pp. 1099–1109. DOI: 10.1109/JMEMS.2005.856676.

[43] Gwi-Eun Song, Kwang-Ho Kim, and Yoon-Pyo Lee. "Simulation and experiments for a phase-change actuator with bistable membrane". In: *Sensors and Actuators A: Physical* 136.2 (May 2007). Micromechanics Section of Sensors and Actuators, pp. 665–672. DOI: 10.1016/j.sna.2006.12.018.

[44] Seunghoon Park and Dooyoung Hah. "Pre-shaped buckled-beam actuators: Theory and experiments". In: *Sensors and Actuators A: Physical* 148.1 (Nov. 2008), pp. 186–192. DOI: 10.1016/j.sna.2008.07.009.

[45] Sahil Oak et al. "Development and Testing of a Multilevel Chevron Actuator-Based Positioning System". In: *Journal of Microelectromechanical Systems* 20.6 (Dec. 2011), pp. 1298–1309. DOI: 10.1109/JMEMS.2011.2167674.

[46] Shin-Wei Huang, Fan-Chi Lin, and Yao-Joe Yang. "A novel single-actuator bistable microdevice with a moment-driven mechanism". In: *Sensors and Actuators A: Physical* 310 (Aug. 2020), p. 111934. DOI: 10.1016/j.sna.2020.111934.

[47] Hussein Hussein, Hossein Fariborzi, and Mohammad I. Younis. "Modeling of Beam Electrothermal Actuators". In: *Journal of Microelectromechanical Systems* 29.6 (Dec. 2020), pp. 1570–1581. DOI: 10.1109/JMEMS.2020.3033477.

[48] Diankun Pan et al. “A novel design and manufacturing method for compliant bistable structure with dissipated energy feature”. In: *Materials & Design* 196 (Nov. 2020), p. 109081. DOI: 10.1016/j.matdes.2020.109081.

[49] Aron Michael and Chee Yee Kwok. “Design criteria for bi-stable behavior in a buckled multi-layered MEMS bridge”. In: *Journal of Micromechanics and Microengineering* 16.10 (Aug. 2006), p. 2034. DOI: 10.1088/0960-1317/16/10/016.

[50] Aron Michael and Chee Yee Kwok. “Thermo-mechanical behavior of buckled multi-layered micro-bridges”. In: *Sensors and Actuators A: Physical* 137.1 (June 2007), pp. 157–168. DOI: 10.1016/j.sna.2007.02.029.

[51] Shengping Mao et al. “A latching bistable microswitch using dual-beam electrothermal actuation”. In: *2010 IEEE 5th International Conference on Nano/Micro Engineered and Molecular Systems*. Xiamen, China, Jan. 2010, pp. 732–735. DOI: 10.1109/NEMS.2010.5592258.

[52] B. Dong et al. “A silicon-nanowire memory driven by optical gradient force induced bistability”. In: *Applied Physics Letters* 107.26 (Dec. 2015), p. 261111. DOI: 10.1063/1.4939114.

[53] Pan Kyu Choi et al. “Electro-Thermally Actuated Non-Volatile Mechanical Memory With CMOS-Level Operation Voltage and Low Contact Resistance”. In: *Journal of Microelectromechanical Systems* 31.1 (Feb. 2022), pp. 87–96. DOI: 10.1109/JMEMS.2021.3120109.

[54] Jixiao Tao et al. “Broadband energy harvesting by using bistable FG-CNTRC plate with integrated piezoelectric layers”. In: *Smart Materials and Structures* 28.9 (Aug. 2019), p. 095021. DOI: 10.1088/1361-665X/ab2dc2.

[55] Mattias Vangbo. “An analytical analysis of a compressed bistable buckled beam”. In: *Sensors and Actuators A: Physical* 69.3 (Sept. 1998), pp. 212–216. DOI: 10.1016/S0924-4247(98)00097-1.

[56] Cho-Chun Wu, Meng-Ju Lin, and Rongshun Chen. “Bistable criterion for mechanically bistable mechanism”. In: *2012 IEEE 25th International Conference on Micro Electro Mechanical Systems (MEMS)*. Paris, France, Jan. 2012, pp. 396–399. DOI: 10.1109/MEMSYS.2012.6170163.

[57] Cho-Chun Wu, Meng-Ju Lin, and Rongshun Chen. “The derivation of a bistable criterion for double V-beam mechanisms”. In: *Journal of Micromechanics and Microengineering* 23.11 (Oct. 2013), p. 115005. DOI: 10.1088/0960-1317/23/11/115005.

[58] B. Camescasse, A. Fernandes, and J. Pouget. “Bistable buckled beam: Elastica modeling and analysis of static actuation”. In: *International Journal of Solids and Structures* 50.19 (Sept. 2013), pp. 2881–2893. DOI: 10.1016/j.ijsolstr.2013.05.005.

[59] Wenzhong Yan, Yunchen Yu, and Ankur Mehta. “Analytical modeling for rapid design of bistable buckled beams”. In: *Theoretical and Applied Mechanics Letters* 9.4 (July 2019), pp. 264–272. DOI: 10.1016/j.taml.2019.04.006.

[60] Ginés Lifante. *Integrated Photonics: Fundamentals*. Wiley, Jan. 2003. DOI: 10.1002/0470861401.

[61] Christopher R. Doerr. "Silicon photonic integration in telecommunications". In: *Frontiers in Physics* 3 (Aug. 2015). DOI: 10.3389/fphy.2015.00037.

[62] Lukas Chrostowski and Michael Hochberg. *Silicon Photonics Design: From Devices to Systems*. Cambridge University Press, Apr. 2015. DOI: 10.1017/cbo9781316084168.

[63] S.M. Sze and Kwok K. Ng. "MOSFETs". In: *Physics of Semiconductor Devices*. First Edition. John Wiley & Sons, Ltd, Apr. 2006. Chap. 6, pp. 293–373. DOI: 10.1002/9780470068328.ch6.

[64] Carlos Errando-Herranz et al. "MEMS for Photonic Integrated Circuits". In: *IEEE Journal of Selected Topics in Quantum Electronics* 26.2 (Mar. 2020), pp. 1–16. DOI: 10.1109/jstqe.2019.2943384.

[65] Jonathan Peltier et al. "High-speed silicon photonic electro-optic Kerr modulation". In: *Photonics Research* 12.1 (Jan. 2024), pp. 51–60. DOI: 10.1364/PRJ.488867.

[66] Niels Quack et al. "Integrated silicon photonic MEMS". In: *Microsystems & Nanoengineering* 9.1 (Mar. 2023), p. 27. DOI: 10.1038/s41378-023-00498-z.

[67] Sangyoon Han et al. "Multicast silicon photonic MEMS switches with gap-adjustable directional couplers". In: *Optics Express* 27.13 (June 2019), pp. 17561–17570. DOI: 10.1364/OE.27.017561.

[68] Sangyoon Han et al. "32 × 32 silicon photonic MEMS switch with gap-adjustable directional couplers fabricated in commercial CMOS foundry". In: *Journal of Optical Microsystems* 1.2 (Mar. 2021), p. 024003. DOI: 10.1117/1.JOM.1.2.024003.

[69] Sangyoon Han et al. "Large-scale silicon photonic switches with movable directional couplers". In: *Optica* 2.4 (Apr. 2015), pp. 370–375. DOI: 10.1364/OPTICA.2.000370.

[70] Hamed Sattari et al. "Bistable silicon photonic MEMS switches". In: *MOEMS and Miniaturized Systems XVIII*. Ed. by Wibool Piyawattanametha, Yong-Hwa Park, and Hans Zappe. Vol. 10931. International Society for Optics and Photonics. SPIE, Mar. 2019, p. 109310D. DOI: 10.1117/12.2507192.

[71] Carlos Errando-Herranz et al. "Low-power microelectromechanically tunable silicon photonic ring resonator add-drop filter". In: *Optics Letters* 40.15 (Aug. 2015), pp. 3556–3559. DOI: 10.1364/OL.40.003556.

[72] Rohit Chatterjee and Chee Wei Wong. "Nanomechanical Proximity Perturbation for Switching in Silicon-Based Directional Couplers for High-Density Photonic Integrated Circuits". In: *Journal of Microelectromechanical Systems* 19.3 (June 2010), pp. 657–662. DOI: 10.1109/JMEMS.2010.2043216.

[73] Xiaojing Wang et al. "Wafer-Level Vacuum Sealing by Transfer Bonding of Silicon Caps for Small Footprint and Ultra-Thin MEMS Packages". In: *Journal of Microelectromechanical Systems* 28.3 (June 2019), pp. 460–471. DOI: 10.1109/JMEMS.2019.2910985.

[74] Horst Theuss and Klaus Pressel. “Three-dimensional integration of MEMS”. In: *Handbook of Silicon Based MEMS Materials and Technologies*. Ed. by Markku Tilli et al. Third Edition. Micro and Nano Technologies. Elsevier, 2020. Chap. 32, pp. 691–706. DOI: 10.1016/B978-0-12-817786-0.00032-3.

[75] Niels Quack et al. “MEMS-Enabled Silicon Photonic Integrated Devices and Circuits”. In: *IEEE Journal of Quantum Electronics* 56.1 (Feb. 2020), pp. 1–10. DOI: 10.1109/JQE.2019.2946841.

[76] J.A. Shaw, C.B. Churchill, and M.A. Iadicola. “Tips and Tricks for Characterizing Shape Memory Alloy Wire: Part 1—Differential Scanning Calorimetry and Basic Phenomena”. In: *Experimental Techniques* 32.5 (Sept. 2008), pp. 55–62. DOI: 10.1111/j.1747-1567.2008.00410.x.

[77] Christoph Haberland, Mahmoud Kadkhodaei, and Mohammad H. Elahinia. “Introduction”. In: *Shape Memory Alloy Actuators: Design, Fabrication, and Experimental Evaluation*. John Wiley & Sons, Ltd, Nov. 2015. Chap. 1, pp. 1–43. DOI: 10.1002/9781118426913.ch1.

[78] David A. Porter, Kenneth E. Easterling, and Mohamed Y. Sherif. *Phase Transformations in Metals and Alloys*. Fourth Edition. Boca Raton: CRC Press, Nov. 2021. DOI: 10.1201/9781003011804.

[79] Shuichi Miyazaki. “Martensitic transformation in TiNi alloys”. In: *Thin Film Shape Memory Alloys: Fundamentals and Device Applications*. Ed. by Shuichi Miyazaki, Yong Qing Fu, and Wei MinEditors Huang. Cambridge University Press, 2009. Chap. 2, pp. 73–87. DOI: 10.1017/CBO9780511635366.

[80] H. E. Karaca et al. “NiTiHf-based shape memory alloys”. In: *Materials Science and Technology* 30.13 (Nov. 2014), pp. 1530–1544. DOI: 10.1179/1743284714Y.0000000598.

[81] Nitin Choudhary and Davinder Kaur. “Shape memory alloy thin films and heterostructures for MEMS applications: A review”. In: *Sensors and Actuators A: Physical* 242 (May 2016), pp. 162–181. DOI: 10.1016/j.sna.2016.02.026.

[82] K. Otsuka and X. Ren. “Physical metallurgy of Ti–Ni-based shape memory alloys”. In: *Progress in Materials Science* 50.5 (July 2005), pp. 511–678. DOI: 10.1016/j.pmatsci.2004.10.001.

[83] A. Ishida, M. Sato, and Z.Y. Gao. “Properties and applications of TiNiCu shape-memory-alloy thin films”. In: *Journal of Alloys and Compounds* 577 (Nov. 2013). SI :ICOMAT2011, S184–S189. DOI: 10.1016/j.jallcom.2011.12.155.

[84] A. David Johnson. “Deposition techniques for TiNi thin film”. In: *Thin Film Shape Memory Alloys: Fundamentals and Device Applications*. Ed. by Shuichi Miyazaki, Yong Qing Fu, and Wei MinEditors Huang. Cambridge University Press, 2009. Chap. 3, pp. 88–109. DOI: 10.1017/CBO9780511635366.

[85] David P. Arnold, Monika Saumer, and Yong-Kyu Yoon. "Additive Processes for Metals". In: *MEMS Materials and Processes Handbook*. Ed. by Reza Ghodssi and Pinyen Lin. Boston, MA: Springer US, Jan. 2011. Chap. 3, pp. 137–191. DOI: 10.1007/978-0-387-47318-5_3.

[86] Y. Q. Fu et al. "Stress and surface morphology evolution". In: *Thin Film Shape Memory Alloys: Fundamentals and Device Applications*. Ed. by Shuichi Miyazaki, Yong Qing Fu, and Wei Min Huang. Cambridge University Press, 2009. Chap. 7, pp. 166–192. DOI: 10.1017/CBO9780511635366.

[87] A Ishida and M Sato. "Thickness effect on shape memory behavior of Ti-50.0at.%Ni thin film". In: *Acta Materialia* 51.18 (Oct. 2003), pp. 5571–5578. DOI: 10.1016/S1359-6454(03)00420-8.

[88] I Jarrige, P Holliger, and P Jonnard. "Diffusion processes in NiTi/Si, NiTi/SiO_2 and NiTi/Si_3N_4 systems under annealing". In: *Thin Solid Films* 458.1 (June 2004), pp. 314–321. DOI: 10.1016/j.tsf.2003.12.039.

[89] Ali S. Turabi et al. "Experimental Characterization of Shape Memory Alloys". In: *Shape Memory Alloy Actuators: Design, Fabrication, and Experimental Evaluation*. John Wiley & Sons, Ltd, Nov. 2015. Chap. 7, pp. 239–277. DOI: 10.1002/9781118426913.ch7.

[90] Maria Rosaria Ricciardi and Vincenza Antonucci. "Experimental characterization of shape memory alloys". In: *Shape Memory Alloy Engineering: For Aerospace, Structural, and Biomedical Applications*. Ed. by Antonio Concilio et al. Second Edition. Boston: Butterworth-Heinemann, Jan. 2021. Chap. 5, pp. 141–164. DOI: 10.1016/B978-0-12-819264-1.00005-4.

[91] I Miccoli et al. "The 100th anniversary of the four-point probe technique: the role of probe geometries in isotropic and anisotropic systems". In: *Journal of Physics: Condensed Matter* 27.22 (May 2015), p. 223201. DOI: 10.1088/0953-8984/27/22/223201.

[92] Ivo Stachiv, Eduardo Alarcon, and Miroslav Lamac. "Shape Memory Alloys and Polymers for MEMS/NEMS Applications: Review on Recent Findings and Challenges in Design, Preparation, and Characterization". In: *Metals* 11.3 (Mar. 2021). DOI: 10.3390/met11030415.

[93] Thomas Mortelmans et al. "Grayscale e-beam lithography: Effects of a delayed development for well-controlled 3D patterning". In: *Microelectronic Engineering* 225 (Feb. 2020), p. 111272. DOI: 10.1016/j.mee.2020.111272.

[94] Sami Franssila and Santeri Tuomikoski. "MEMS lithography". In: *Handbook of Silicon Based MEMS Materials and Technologies*. Ed. by Markku Tilli et al. Third Edition. Micro and Nano Technologies. Elsevier, 2020. Chap. 15, pp. 399–416. DOI: 10.1016/B978-0-12-817786-0.00015-3.

[95] Helmut Kohl and Ludwig Reimer. "Wave Optics of Electrons". In: *Transmission Electron Microscopy: Physics of Image Formation*. Fifth Edition. New York, NY: Springer New York, July 2008. Chap. 3, pp. 43–74. DOI: 10.1007/978-0-387-40093-8_3.

[96] Daniel R. Hines et al. “MEMS Lithography and Micromachining Techniques”. In: *MEMS Materials and Processes Handbook*. Ed. by Reza Ghodssi and Pinyen Lin. Boston, MA: Springer US, Jan. 2011. Chap. 9, pp. 667–753. DOI: 10.1007/978-0-387-47318-5_9.

[97] Zheng Cui. “Electron Beam Lithography”. In: *Nanofabrication: Principles, Capabilities and Limits*. Third Edition. Cham: Springer International Publishing, July 2024. Chap. 3, pp. 83–139. DOI: 10.1007/978-3-031-62546-6_3.

[98] A E Grigorescu and C W Hagen. “Resists for sub-20-nm electron beam lithography with a focus on HSQ: state of the art”. In: *Nanotechnology* 20.29 (July 2009), p. 292001. DOI: 10.1088/0957-4484/20/29/292001.

[99] Bruce W. Smith and Kazuaki Suzuki. *Microlithography: Science and Technology*. Third Edition. Boca Raton: CRC Press, May 2020. DOI: 10.1201/9781315117171.

[100] Ervin B. Podgoršak. “Coulomb Scattering”. In: *Radiation Physics for Medical Physicists*. Third Edition. Cham: Springer International Publishing, Nov. 2016. Chap. 2, pp. 79–142. DOI: 10.1007/978-3-319-25382-4_2.

[101] Stephanie Nicole Bonvicini et al. “Formation of Au, Pt, and bimetallic Au–Pt nanostructures from thermal dewetting of single-layer or bilayer thin films”. In: *Nanotechnology* 33.23 (Mar. 2022), p. 235604. DOI: 10.1088/1361-6528/ac5a83.

[102] Min Zhao et al. “Technology of alignment mark in electron beam lithography”. In: *7th International Symposium on Advanced Optical Manufacturing and Testing Technologies: Smart Structures and Materials for Manufacturing and Testing*. Ed. by Xiangang Luo and Harald Giessen. Vol. 9285. International Society for Optics and Photonics. Harbin, China: SPIE, Aug. 2014, p. 92850C. DOI: 10.1117/12.2068112.

[103] Faiz Rahman et al. “A review of polymethyl methacrylate (PMMA) as a versatile lithographic resist – With emphasis on UV exposure”. In: *Microelectronic Engineering* 224 (Mar. 2020), p. 111238. DOI: 10.1016/j.mee.2020.111238.

[104] Haifang Yang et al. “Electron beam lithography of HSQ/PMMA bilayer resists for negative tone lift-off process”. In: *Microelectronic Engineering* 85.5 (Jan. 2008). Proceedings of the Micro- and Nano-Engineering 2007 Conference, pp. 814–817. DOI: 10.1016/j.mee.2008.01.006.

[105] V M Sundaram and S-B Wen. “An easy method to perform e-beam negative tone lift-off fabrication on dielectric material with a sandwiched conducting polymer layer”. In: *Journal of Micromechanics and Microengineering* 21.6 (May 2011), p. 065021. DOI: 10.1088/0960-1317/21/6/065021.

[106] Stephen Thoms and Douglas S. Macintyre. “Long nanoscale gaps on III–V substrates by electron beam lithography”. In: *Journal of Vacuum Science & Technology B* 30.6 (Nov. 2012), 06F305. DOI: 10.1116/1.4766881.

[107] Huigao Duan et al. “Free-standing sub-10 nm nanostencils for the definition of gaps in plasmonic antennas”. In: *Nanotechnology* 24.18 (Apr. 2013), p. 185301. DOI: 10.1088/0957-4484/24/18/185301.

[108] Marcus Rommel et al. “Sub-10nm resolution after lift-off using HSQ/PMMA double layer resist”. In: *Microelectronic Engineering* 110 (Mar. 2013), pp. 123–125. DOI: 10.1016/j.mee.2013.02.101.

[109] Jintang Huang et al. “Fabrication of ultrahigh resolution metal nanowires and nanodots through EUV interference lithography”. In: *Microelectronic Engineering* 141 (Jan. 2015). Micro/Nano Fabrication 2014, pp. 32–36. DOI: 10.1016/j.mee.2015.01.016.

[110] Srinivas Tadigadapa and Franz Lärmer. “Dry Etching for Micromachining Applications”. In: *MEMS Materials and Processes Handbook*. Ed. by Reza Ghodssi and Pinyen Lin. Boston, MA: Springer US, Jan. 2011. Chap. 7, pp. 403–456. DOI: 10.1007/978-0-387-47318-5_7.

[111] Stefan Tinck et al. “Concurrent effects of wafer temperature and oxygen fraction on cryogenic silicon etching with SF_6/O_2 plasmas”. In: *Plasma Processes and Polymers* 14.9 (Sept. 2017), p. 1700018. DOI: 10.1002/ppap.201700018.

[112] Qi Chen et al. “Isotropic etch for SiO_2 microcantilever release with ICP system”. In: *Microelectronic Engineering* 85.3 (Mar. 2008), pp. 500–507. DOI: 10.1016/j.mee.2007.09.004.

[113] Ü. Sökmen et al. “Capabilities of ICP-RIE cryogenic dry etching of silicon: review of exemplary microstructures”. In: *Journal of Micromechanics and Microengineering* 19.10 (Sept. 2009), p. 105005. DOI: 10.1088/0960-1317/19/10/105005.

[114] Franz Laermer et al. “Deep reactive ion etching”. In: *Handbook of Silicon Based MEMS Materials and Technologies*. Ed. by Markku Tilli et al. Third Edition. Micro and Nano Technologies. Elsevier, 2020. Chap. 16, pp. 417–446. DOI: 10.1016/B978-0-12-817786-0.00016-5.

[115] N Chekurov et al. “Atomic layer deposition enhanced rapid dry fabrication of micromechanical devices with cryogenic deep reactive ion etching”. In: *Journal of Micromechanics and Microengineering* 17.8 (July 2007), p. 1731. DOI: 10.1088/0960-1317/17/8/041.

[116] Martin Drost et al. “Etch mechanism of an Al_2O_3 hard mask in the Bosch process”. In: *Micro and Nano Engineering* 14 (Apr. 2022), p. 100102. DOI: 10.1016/j.mne.2021.100102.

[117] David W. Burns. “MEMS Wet-Etch Processes and Procedures”. In: *MEMS Materials and Processes Handbook*. Ed. by Reza Ghodssi and Pinyen Lin. Boston, MA: Springer US, Jan. 2011. Chap. 8, pp. 457–665. DOI: 10.1007/978-0-387-47318-5_8.

[118] Patrick Echlin. “Sample Dehydration”. In: *Handbook of Sample Preparation for Scanning Electron Microscopy and X-Ray Microanalysis*. Boston, MA: Springer US, Jan. 2009. Chap. 7, pp. 97–136. DOI: 10.1007/978-0-387-85731-2_7.

[119] David A. B. Miller. “Device Requirements for Optical Interconnects to Silicon Chips”. In: *Proceedings of the IEEE* 97.7 (July 2009), pp. 1166–1185. DOI: 10.1109/JPROC.2009.2014298.

[120] Emanuele Pelucchi et al. "The potential and global outlook of integrated photonics for quantum technologies". In: *Nature Reviews Physics* 4.3 (Mar. 2022), pp. 194–208. DOI: 10.1038/s42254-021-00398-z.

[121] Graham T. Reed and Andrew P. Knights. *Silicon Photonics: An Introduction*. John Wiley & Sons, Ltd, Jan. 2004. DOI: 10.1002/0470014180.

[122] Katsunari Okamoto. "Planar Optical Waveguides". In: *Fundamentals of Optical Waveguides*. Third Edition. Elsevier, Jan. 2022. Chap. 2, pp. 13–65. DOI: 10.1016/b978-0-12-815601-8.50002-1.

[123] Pablo Sanchis et al. "Design of Silicon-Based Slot Waveguide Configurations for Optimum Nonlinear Performance". In: *Journal of Lightwave Technology* 25.5 (May 2007), pp. 1298–1305. DOI: 10.1109/JLT.2007.893909.

[124] Xinru Ji et al. "Compact, spatial-mode-interaction-free, ultralow-loss, nonlinear photonic integrated circuits". In: *Communications Physics* 5.1 (Apr. 2022). DOI: 10.1038/s42005-022-00851-0.

[125] Takeshi Fujisawa et al. "Low-loss, compact, and fabrication-tolerant Si-wire 90° waveguide bend using clothoid and normal curves for large scale photonic integrated circuits". In: *Optics Express* 25.8 (Apr. 2017), pp. 9150–9159. DOI: 10.1364/oe.25.009150.

[126] Florian Vogelbacher et al. "Analysis of silicon nitride partial Euler waveguide bends". In: *Optics Express* 27.22 (Oct. 2019), pp. 31394–31406. DOI: 10.1364/OE.27.031394.

[127] Dominik Gerhard Rabus and Cinzia Sada. *Integrated Ring Resonators: A Compendium*. Second edition. Springer International Publishing, Dec. 2020. DOI: 10.1007/978-3-030-60131-7.

[128] Takumi Nagai and Kazuhiro Hane. "Silicon photonic microelectromechanical switch using lateral adiabatic waveguide couplers". In: *Optics Express* 26.26 (Dec. 2018), pp. 33906–33917. DOI: 10.1364/oe.26.033906.

[129] Heba Tamazin et al. "Ultra-broadband Compact Adiabatic Coupler in Silicon-on-Insulator for Joint Operation in the C- and O-Bands". In: *Conference on Lasers and Electro-Optics*. Optica Publishing Group, May 2018. DOI: 10.1364/CLEO_SI.2018.STh4B.4.

[130] Jinghui Zou. "Silicon based Grating Coupler for Fiber-chip Coupling (in Chinese)". PhD thesis. Huazhong University of Science and Technology, May 2017.

[131] Robert Halir et al. "Waveguide grating coupler with subwavelength microstructures". In: *Optics Letters* 34.9 (May 2009), pp. 1408–1410. DOI: 10.1364/ol.34.001408.

[132] R. Halir et al. "Continuously apodized fiber-to-chip surface grating coupler with refractive index engineered subwavelength structure". In: *Optics Letters* 35.19 (Oct. 2010), pp. 3243–3245. DOI: 10.1364/ol.35.003243.

[133] Mikael Antelius, Kristinn B. Gylfason, and Hans Sohlström. "An apodized SOI waveguide-to-fiber surface grating coupler for single lithography silicon photonics". In: *Optics Express* 19.4 (Feb. 2011), pp. 3592–3598. DOI: 10.1364/oe.19.003592.

[134] Daniel Benedikovic et al. “High-efficiency single etch step apodized surface grating coupler using subwavelength structure”. In: *Laser & Photonics Reviews* 8.6 (2014), pp. L93–L97. DOI: 10.1002/lpor.201400113.

[135] Lirong Cheng et al. “Grating Couplers on Silicon Photonics: Design Principles, Emerging Trends and Practical Issues”. In: *Micromachines* 11.7 (July 2020). DOI: 10.3390/mi11070666.

[136] Larry L. Howell and James N. Leonard. “Optimal loading conditions for non-linear deflections”. In: *International Journal of Non-Linear Mechanics* 32.3 (1997), pp. 505–514. DOI: 10.1016/S0020-7462(96)00069-8.

[137] Chris Kimball and Lung-Wen Tsai. “Modeling of Flexural Beams Subjected to Arbitrary End Loads”. In: *Journal of Mechanical Design* 124.2 (May 2002), pp. 223–235. DOI: 10.1115/1.1455031.

[138] Aimei Zhang and Guimin Chen. “A Comprehensive Elliptic Integral Solution to the Large Deflection Problems of Thin Beams in Compliant Mechanisms”. In: *Journal of Mechanisms and Robotics* 5.2 (Mar. 2013), p. 021006. DOI: 10.1115/1.4023558.

[139] James M. Gere and Barry J. Goodno. *Mechanics of materials*. eng. Seventh Edition. Includes bibliographical references (p. 935-942) and indexes. Clifton Park, NY: Cengage Learning, 2009.

[140] George Gerald Stoney and Charles Algernon Parsons. “The tension of metallic films deposited by electrolysis”. In: *Proceedings of the Royal Society of London. Series A* 82.553 (1909), pp. 172–175. DOI: 10.1098/rspa.1909.0021.

[141] Dario Grochla et al. “Si micro-cantilever sensor chips for space-resolved stress measurements in physical and plasma-enhanced chemical vapour deposition”. In: *Sensors and Actuators A: Physical* 270 (2018), pp. 271–277. DOI: 10.1016/j.sna.2017.12.050.

[142] Grégory Abadias et al. “Review Article: Stress in thin films and coatings: Current status, challenges, and prospects”. In: *Journal of Vacuum Science & Technology A* 36.2 (Mar. 2018), p. 020801. DOI: 10.1116/1.5011790.

[143] Michael Huff. “Review Paper: Residual Stresses in Deposited Thin-Film Material Layers for Micro- and Nano-Systems Manufacturing”. In: *Micromachines* 13.12 (2022). DOI: 10.3390/mi13122084.

[144] Theodore L. Bergman et al. *Fundamentals of Heat and Mass Transfer*. John Wiley & Sons, 2011.

[145] Gowtham Arivanandhan et al. “Temperature Homogenization of Co-Integrated Shape Memory–Silicon Bimorph Actuators”. In: *Proceedings* 64.1 (2020). DOI: 10.3390/IeCAT2020-08501.

[146] Gowtham Arivanandhan et al. “Power Optimization of TiNiHf/Si Shape Memory Microactuators”. In: *Actuators* 12.2 (2023). DOI: 10.3390/act12020082.

[147] Richard G. Budynas and Ali M. Sadegh. *Roark’s Formulas for Stress and Strain*. en. Ninth Edition. New York: McGraw-Hill Education, 2020.

[148] Ronald F. Gibson. *Principles of Composite Material Mechanics*. Fourth Edition. CRC Press, Feb. 2016. DOI: 10.1201/b19626.

[149] K. Ikuta and H. Shimizu. "Two dimensional mathematical model of shape memory alloy and intelligent SMA-CAD". In: *[1993] Proceedings IEEE Micro Electro Mechanical Systems*. IEEE, 1993, pp. 87–92. DOI: 10.1109/memsys.1993.296958.

[150] P. Sedlák et al. "Thermomechanical model for NiTi-based shape memory alloys including R-phase and material anisotropy under multi-axial loadings". In: *International Journal of Plasticity* 39 (2012), pp. 132–151. DOI: https://doi.org/10.1016/j.ijplas.2012.06.008.

[151] Yongqing Fu et al. "Stress and surface morphology of TiNiCu thin films: effect of annealing temperature". In: *Surface and Coatings Technology* 198.1 (Aug. 2005), pp. 389–394. DOI: 10.1016/j.surfcoat.2004.10.107.

[152] Simon M. Sze, Yiming Li, and Kwok K. Ng. *Physics of Semiconductor Devices*. Fourth Edition. John Wiley & Sons, Mar. 2021.

[153] H. R. Shanks et al. "Thermal Conductivity of Silicon from 300 to 1400°K". In: *Physical Review* 130.5 (June 1963), pp. 1743–1748. DOI: 10.1103/physrev.130.1743.

[154] P.G. Neudeck, R.S. Okojie, and Liang-Yu Chen. "High-temperature electronics - a role for wide bandgap semiconductors?" In: *Proceedings of the IEEE* 90.6 (June 2002), pp. 1065–1076. DOI: 10.1109/jproc.2002.1021571.

[155] Kangjie Chu et al. "Grain size effect on the temperature-dependence of elastic modulus of nanocrystalline NiTi". In: *Journal of Alloys and Compounds* 934 (Feb. 2023), p. 167907. DOI: 10.1016/j.jallcom.2022.167907.

[156] Minglu Xia, Pan Liu, and Qingping Sun. "Grain Size Effects on Young's Modulus and Hardness of Nanocrystalline NiTi Shape Memory Alloy". In: *Advances in Shape Memory Materials*. Ed. by Qingping Sun et al. Vol. 73. Advanced Structured Materials. Cham: Springer International Publishing, Mar. 2017. Chap. 15, pp. 203–210. DOI: 10.1007/978-3-319-53306-3_15.

[157] Dennis König et al. "Micro- to Nanostructured Devices for the Characterization of Scaling Effects in Shape-Memory Thin Films". In: *Journal of Microelectromechanical Systems* 19.5 (Oct. 2010), pp. 1264–1269. DOI: 10.1109/JMEMS.2010.2067441.

[158] F. Yang et al. "Couple stress based strain gradient theory for elasticity". In: *International Journal of Solids and Structures* 39.10 (May 2002), pp. 2731–2743. DOI: 10.1016/S0020-7683(02)00152-X.

[159] D.C.C. Lam et al. "Experiments and theory in strain gradient elasticity". In: *Journal of the Mechanics and Physics of Solids* 51.8 (Aug. 2003), pp. 1477–1508. DOI: 10.1016/S0022-5096(03)00053-X.

[160] S K Park and X-L Gao. "Bernoulli–Euler beam model based on a modified couple stress theory". In: *Journal of Micromechanics and Microengineering* 16.11 (Sept. 2006), p. 2355. DOI: 10.1088/0960-1317/16/11/015.

[161] W. Xia, L. Wang, and L. Yin. "Nonlinear non-classical microscale beams: Static bending, postbuckling and free vibration". In: *International Journal of Engineering Science* 48.12 (Dec. 2010), pp. 2044–2053. DOI: 10.1016/j.ijengsci.2010.04.010.

[162] Huu-Tai Thai et al. "A review of continuum mechanics models for size-dependent analysis of beams and plates". In: *Composite Structures* 177 (Oct. 2017), pp. 196–219. DOI: 10.1016/j.compstruct.2017.06.040.

[163] Zou-Qing Tan and Yang-Chun Chen. "Size-dependent electro-thermo-mechanical analysis of multilayer cantilever microactuators by Joule heating using the modified couple stress theory". In: *Composites Part B: Engineering* 161 (Mar. 2019), pp. 183–189. DOI: 10.1016/j.compositesb.2018.10.067.

[164] Mir Abbas Roudbari et al. "A review of size-dependent continuum mechanics models for micro- and nano-structures". In: *Thin-Walled Structures* 170 (Jan. 2022), p. 108562. DOI: 10.1016/j.tws.2021.108562.

[165] Jae-Hoon Choi, Wael Zaki, and Gi-Dong Sim. "Size-dependent constitutive model for shape memory alloys based on couple stress elastoplasticity". In: *Applied Mathematical Modelling* 118 (June 2023), pp. 641–664. DOI: 10.1016/j.apm.2023.01.038.

[166] Kuo-Shen Chen and Kuang-Shun Ou. "MEMS residual stress characterization: methodology and perspective". In: *Handbook of Silicon Based MEMS Materials and Technologies*. Ed. by Markku Tilli et al. Third Edition. Micro and Nano Technologies. Elsevier, 2020. Chap. 39, pp. 787–801. DOI: 10.1016/B978-0-12-817786-0.00039-6.

[167] Joachim Piprek. "Optical Waves". In: *Semiconductor Optoelectronic Devices: Introduction to Physics and Simulation*. Boston: Academic Press, 2003. Chap. 4, pp. 83–120. DOI: 10.1016/B978-0-08-046978-2.50029-6.

[168] Y. Laghla and E. Scheid. "Optical study of undoped, B or P-doped polysilicon". In: *Thin Solid Films* 306.1 (Aug. 1997), pp. 67–73. DOI: 10.1016/S0040-6090(97)00247-2.

[169] Mark Auslender and Shlomo Hava. "Doped n-Type Silicon (n-Si)". In: *Handbook of Optical Constants of Solids*. Ed. by Edward D. Palik. Vol. 3. Burlington: Academic Press, 1997. Chap. 6, pp. 155–186. DOI: 10.1016/B978-012544415-6.50102-3.

[170] S. Basu, B. J. Lee, and Z. M. Zhang. "Infrared Radiative Properties of Heavily Doped Silicon at Room Temperature". In: *Journal of Heat Transfer* 132.2 (Nov. 2009), p. 023301. DOI: 10.1115/1.4000171.

[171] "Thermo-Optic Coefficients". In: *Handbook of Optical Constants of Solids*. Ed. by Edward D. Palik. Vol. 5. Burlington: Academic Press, 1997. Chap. 3, pp. 115–261. DOI: 10.1016/B978-012544415-6.50150-3.

[172] Dan-Xia Xu et al. "Empirical model for the temperature dependence of silicon refractive index from O to C band based on waveguide measurements". In: *Optics Express* 27.19 (Sept. 2019), pp. 27229–27241. DOI: 10.1364/OE.27.027229.

[173] Hamed Sattari et al. "Compact broadband suspended silicon photonic directional coupler". In: *Optics Letters* 45.11 (June 2020), pp. 2997–3000. DOI: 10.1364/OL.394470.

A. Partial Euler Bend

For an arbitrary curve s in planar Cartesian coordinate system, as depicted in Fig. A.1, relation between curvature $\varphi(s)$ at point $(x(s), y(s))$ and bend angle $\phi(s)$, which is marked by the angle between $x-$axis and tangential line of curve at this point, is

$$\varphi(s) = \frac{\mathrm{d}\phi(s)}{\mathrm{d}s}, \tag{A.1}$$

$$\phi(s) = \phi(0) + \int_0^s \varphi(u)\,\mathrm{d}u. \tag{A.2}$$

Therefore, coordinates $(x(s), y(s))$ are

$$\begin{aligned} x(s) &= x(0) + \int_0^s \cos(\phi(u))\,\mathrm{d}u, \\ y(s) &= y(0) + \int_0^s \sin(\phi(u))\,\mathrm{d}u, \end{aligned} \tag{A.3}$$

respectively.

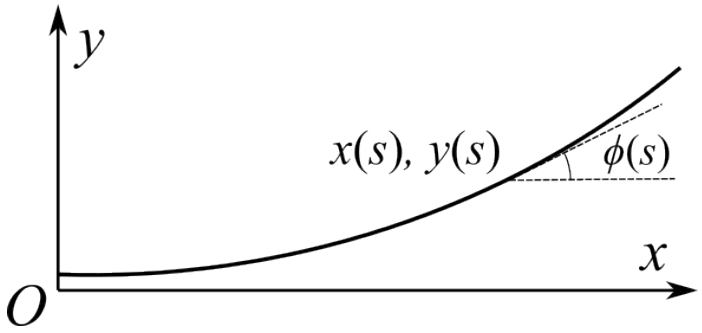


Figure A.1.: An arbitrary curve in planar Cartesian coordinate system. The angle between $x-$axis and tangential line of curve at point $(x(s), y(s))$ is denoted as $\phi(s)$.

Because the curvature of Euler bend is linear to its length, i.e.,

$$\varphi(s) = \alpha s, \tag{A.4}$$

where α is a geometrical parameter, by integrating Eq. A.2–Eq. A.4, the coordinates in Eq. A.3 can be reconstructed in the form of Fresnel integrals

$$\begin{aligned} x(s) &= \int_0^s \cos\left(\frac{\alpha}{2} \cdot u^2\right)\mathrm{d}u, \\ y(s) &= \int_0^s \sin\left(\frac{\alpha}{2} \cdot u^2\right)\mathrm{d}u. \end{aligned} \tag{A.5}$$

For simplicity, here the initial coordinates $x(0), y(0), \phi(0)$ are all set to zero:

$$x(0) = 0; \quad y(0) = 0; \quad \phi(0) = 0. \tag{A.6}$$

In Section 3.2.3, it has been explained that a partial Euler bend is composed by two identical Euler bends and one arc bend sharing the same curvature at connecting points, as shown in Fig. A.2. Necessary parameters to describe the partial Euler bend are also marked out, which include the effective and minimum radii R_{eff} and R_{min}, total bend angle Φ and angle ratio p. Coordinates of connecting point of the first Euler bend and arc bend is (x_p, y_p) and the final point of total bend is (x_L, y_L). In design of PICs, usually values of variables Φ, p, R_{eff} or R_{min} are preset, in the following, solutions of other relevant geometrical parameters will be presented based on geometrical relations.

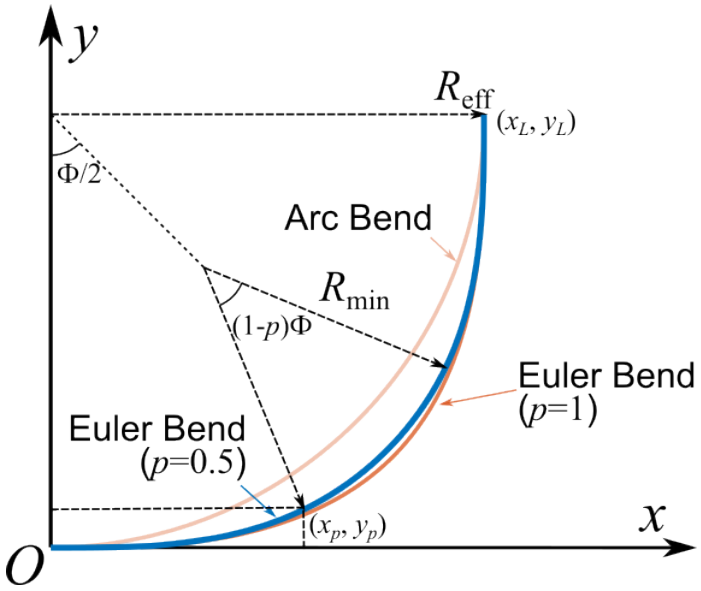


Figure A.2.: Partial Euler bend and necessary defining parameters. An arc bend and Euler bend with $p = 1$ are also drawn for comparison.

At point (x_p, y_p) Based on Eq. A.1 and Eq. A.4, bend angle of Euler bend ϕ_p and the minimum radius, which is also the radius at connecting point of Euler bend and arc bend, are

$$\phi_p = p\Phi = \frac{\alpha s_p^2}{2}, \tag{A.7}$$

$$R_{\text{min}} = \frac{1}{\alpha s_p}, \tag{A.8}$$

respectively.

At point (x_L, y_L) Total length of partial Euler bend L is

$$L = 2s_p + (1 - p)\Phi R_{\min}. \tag{A.9}$$

From Fig. A.2, y_L can be expressed by R_{eff} and Φ

$$y_L = R_{\text{eff}}(1 - \cos\Phi). \tag{A.10}$$

In the meantime, based on Eq. A.2 and Eq. A.3, y_L can also be obtained by integrating along different sections of bends

$$y_L = \int_0^L \sin(\phi(u))\,\mathrm{d}u \tag{A.11}$$

where

$$\phi(u) = \begin{cases} \dfrac{\alpha}{2}u^2, & \text{if} \quad 0 \leqslant u < s_p \\ \dfrac{u}{R_{\min}} + C_1, & \text{if} \quad s_p \leqslant u < L - s_p \\ \dfrac{\alpha}{2}u^2 + C_2, & \text{if} \quad L - s_p \leqslant u < L \end{cases} \tag{A.12}$$

Integration constants C_1, C_2 can be determined by the consecutive conditions of this piecewise function

$$\begin{aligned} \frac{\alpha}{2}s_p^2 &= \frac{s_p}{R_{\min}} + C_1 \\ \frac{L - s_p}{R_{\min}} + C_1 &= \frac{\alpha}{2}(L - s_p)^2 + C_2 \end{aligned} \tag{A.13}$$

Combining Eq. A.7–Eq. A.13 can give the solutions of $\alpha, s_p, R_{\min}$ and L.

Acknowledgment

I can still clearly remember the first time that I entered Germany in 2016, which was two days before the Christmas Eve. I planned a trip from Lyon to Köln and changed the train at Mannheim Hbf in a dark at cold evening. At that time, I would never expect that after four years, I would change the train at this station again, go to a city 70 km away and spend almost five years there to pursue PhD. Recalling this academic journey full of twists and turns from the very beginning, I still cannot believe that I have reached the destination, finally.

The trip of pursuing PhD is overwhelmed with all the emotions that you can image: the exciting exploration of new field of knowledge, the surprising results from experiments, the sad email of rejection and the tired drafting of a paper. The successful completion of this dissertation would not have been possible without the support and guidance of many individuals and groups, to whom I owe my deepest gratitude.

First and foremost, I would like to express my sincere appreciation to my supervisor, Prof. Dr. Manfred Kohl, for his invaluable guidance, insightful feedback, and unwavering support throughout my research. His expertise and rigorous academic standards have profoundly shaped this work, and I am deeply grateful for his mentorship. I would also extend my gratitude to my co-referee, Prof. Dr.-Ing. Alfred Ludwig, for his constructive critiques and efforts of refining this dissertation.

My sincere thanks go to my partners of the *KOMMMA* project. Elaheh Akbarnejad from Ruhr University Bochum (RUB) and Sabrina Curtis from Kiel University (CAU) provided me with excellent Ti-Ni-Cu and Ti-Ni-Hf films during the second and first phase of the project. The high-quality materials were the fundamentals of my fabrication work. Definitely I can't overlook the supports from my partners working on simulations — Muhammad Babar Shamim and Marian Sielenkämper from CAU. Their sophisticated numeric models could always quickly verify the conceived structural designs.

My research has also been supported by lots of colleagues at IMT from different groups, to whom I would like to express my heartfelt appreciation. The first thanks go to the Ebeam group: Andreas Bacher, Marie-Kristin Gamer and Dr. Judith Hohmann. We have closely worked together to achieve high-fidelity exposure results, and they were always helpful and ready to answer my needs. Alban Muslija shared his expertise on SEM and RIE. Anja Ehrhardt and Birgit Hübner offered me with films prepared by evaporation and ALD. Though the direct laser writing technique is not included in my final fabrication processes, I still appreciate the efforts that the DLW team, especially Stefan Hengsbach and Samuel Bergdolt, has devoted to my exposure requests. Heike Fornasier and Dr. Uwe Köhler have

supported me with the cleanroom facilities, SiO_2 wet etching and wafer dicing. Alexandra Mortiz and her group helped me with the fabrication of components for building the setup for optic transmission test. The pneumatic parts of the setup were supported by Marco Heiler. The helpful institute IT administrators, Stefan Vautrelle and Manuel Frisman, offered firm supports on the simulation programs.

I am deeply grateful to my colleagues, Gowtham Arivanandhan, Vincent Gottwald and Joel Joseph, for their always fast responses to my needs of etching, insightful discussions on the experiments, assistance with laser cutting and FemtoTools. The sincere gratitude extends to Alexander Kotz from IPQ for his detailed and patient guidance to the optic transmission tests at the initial stage of my work, as well as the fiber repairing and assistance on the FDTD simulations.

My sincere thanks also go to all SMD group members for all their helps in the past five years. I would like to thank Danni Li, Lena Seigner, Carina Ludwig and Xi Chen. It is my fortune that I can be part of this excellent group.

I gratefully acknowledge DFG for providing funding to the priority program "Cooperative Multistage Multistable Microactuator Systems (KOMMMA)". I can have the unforgettable opportunity to pursue PhD in Germany and meet people from various backgrounds across the whole Germany.

Finally, I owe my deepest gratitude to my family for their unconditional support and encouragement throughout this journey. This appreciation extends to my girlfriend, Jiayu Chen. The geographic distance between Karlsruhe and Delft didn’t separate our hearts. I will always remember our treasured memories on the snowy Alpes, in the tranquil Black Forest and at the sunny Sicily.

Karlsruhe, June 2025

Zixiong Li

Publications

Drafted

Zixiong Li et al. "A SMA multilayer microactuator with out-of-plane bistability". In: *50th International Micro and Nano Engineering Conference.* Montpellier, France, Sep. 2024.

Zixiong Li et al. "Bistable Actuation Based on TiNiHf/Si Shape Memory Nanoactuators". In: *MikroSystemTechnik Kongress 2023; Kongress.* VDE. Dresden, Germany, Oct. 2023, pp. 93–97. (Link)

Zixiong Li et al. "Design and Fabrication of a Photonic Waveguide Switch Based on a VO_2 Film". In: *47th International Micro and Nano Engineering Conference.* Turin, Italy, Sep. 2021.

Zixiong Li et al. "Development of Co-Integrated Shape Memory Actuators for Silicon Micro- and Nanomechanics". In: *ACTUATOR; International Conference and Exhibition on New Actuator Systems and Applications 2021.* VDE. Feb. 2021, pp. 1–2. (Link)

Contributed to

Manfred Kohl, …, **Zixiong Li** et al. "Active bi- and multistability in cooperative microactuator systems". In: *Sensors and Actuators Reports* 9 (June 2025), p. 100338. (DOI: 10.1016/j.snr.2025.100338)

Sabrina M. Curtis, …, **Zixiong Li** et al. "TiNiHf/SiO_2/Si shape memory film composites for bi-directional micro actuation". In: *International Journal of Smart and Nano Materials* 13.2 (May 2022), pp. 293–314. (DOI: 10.1080/19475411.2022.2071352)