Monolithically Integrated MEMS Resonators and Oscillators in Standard IC Technology

by

Bichoy Waguih Azmy Bahr

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Electrical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2016

Author . . . .
Department of Electrical Engineering and Computer Science
May 18, 2016

Certified by . . . .
Dana Weinstein
Associate Professor
Thesis Supervisor

Certified by . . . .
Luca Daniel
Professor
Thesis Supervisor

Accepted by . . . .
Leslie A. Kolodziejski
Chairman, Department Committee on Graduate Theses

Monolithically Integrated MEMS Resonators and Oscillators in Standard IC Technology

by

Bichoy Waguih Azmy Bahr

Submitted to the Department of Electrical Engineering and Computer Science
on May 18, 2016, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Electrical Engineering

Abstract

Frequency sources and high quality filters based on mechanical resonators are essential building blocks for communication systems as well as analog and digital electronics. Driven by the continuous demand for reduction in power, size and overall cost, monolithic integration of mechanical resonators in standard integrated circuit (IC) technology has been the focus of multiple research efforts. Micro-Electro-Mechanical (MEM) resonators offer an ultimate solution, with 100× higher quality factors and $10^4\times$ smaller footprint, when compared to on-chip LC tank circuits.

A new class of truly solid-state, monolithically integrated, GHz-frequencies CMOS-MEMS resonators is presented. No post-processing or special packaging of any kind is required beyond the standard CMOS process. Resonant body transistor (RBT) is constructed by using active field-effect-transistor (FET) sensing.

A phononic crystal (PnC) implemented in the CMOS back-end-of-line (BEOL) layers along with the bulk wafer are used to create a phononic waveguide. The latter confines acoustic vibrations in the CMOS front-end-of-line (FEOL) layers. Operator-theoretic analysis for these waveguides is presented in explicit analogy to quantum mechanics and photonic waveguides; with a study of perturbation theory, coupled-mode theory and adiabatic theorem. Superior energy confinement is achieved, allowing record high $Q \sim 14,800$ and $f_\circ \cdot Q \sim 4.85 \times 10^{13}$ for CMOS-RBTs at 3 GHz. Simulation, modeling, optimization, prototyping and testing of these resonators is presented. RBTs in FinFET technologies are also exploerd, for resonance frequencies up to 33 GHz.

The thesis also explores the integration of Lamb-mode resonators in standard GaN monolithic-microwave-IC (MMIC) process. The first monolithic 1GHz MEMS-based oscillator in standard GaN MMIC technology is demonstrated, together with monolithic lattice and ladder filters. This allows for complete RF front-ends in GaN MMIC technology.

Thesis Supervisor: Dana Weinstein
Title: Associate Professor

Thesis Supervisor: Luca Daniel
Title: Professor

Acknowledgments

First and foremost, glory and thanks goes to my personal Lord and Savior, Jesus Christ; for whom I owe my very life. I thank Him for His grace that supported and guided me through my entire life: “And the Lord was with Joseph, and he was a prosperous man;” Genesis 39:2a.

During my study at MIT, I worked with many amazing people who deserves my thanks and gratitude. First of all, I am deeply indebted to my PhD advisor Prof. Dana Weinstein for her endless patience, encouragement, and mentoring throughout the years of my PhD. Her continued advice, direction, and guidance helped me to navigate the most intricate challenges. Prof. Weinstein consistently ignited my inspiration and encouraged me to develop new ideas. She continuously believed in me and made sure that I always have the best equipment and resources to succeed and thrive. Her sincere mentorship and support were crucial for starting my career. My extreme gratitude also goes to Prof. Luca Daniel, my PhD co-advisor. His extended support, guidance, and mentorship were invaluable to my success. Prof. Daniel always helped me to look at my work from a new perspective, which was very inspiring. I was exceedingly honored and blessed by working with both Prof. Weinstein and Prof. Daniel.

I am also grateful for Prof. Dimitri Antoniadis for being a member of my thesis committee. His feedback and comments were particularly useful and contributed positively to my work.

I would also like to thank my colleagues in the HybridMEMS Group for the exciting discussions and useful collaborations, namely: Wentao Wang, Laura Popa, Siping Wang, Subramanian Sundaram, Wen-Chien Chen, Bethany Kroese, Matthew Storey, and Tianyi Zeng.

In particular, I would like to extend my gratitude to Wentao Wang for the great many enlightening discussions we had. Wentao was very generous with his time: he introduced me to new concepts and helped me to deepen my understanding of several others. I am also grateful to Siping Wang for the inspiring physics discussion. I greatly enjoyed Wentao's and Siping's exceptional work ethics and camaraderie.

Also, I would like to thank Laura Popa for introducing me to the GaN MMIC technology. The fruitful collaboration we had led to the monolithic GaN MEMS oscillators project. Her invaluable contribution made such a project possible: she designed, optimized and executed the GaN MEMS/MMIC fabrication process, along with the GaN MEMS resonators. I am also grateful to Ujwal Radhakrishna for his extensive help in the development and calibration of the HEMT models.

A special thanks goes to my collaborators from Politecnico di Milano. I am very grateful to Prof. Paolo Maffezzoni for introducing me to the reduced order modeling of oscillators and his unmatched contribution to the coupled oscillators project. I would also like to thank Prof. Raffaele Ardito and Luca D'Alessandro for their tremendous efforts and contributions to the shape optimization of phononic crystals project.

My sincere gratitude also goes to Prof. Steven Johnson for introducing me to the theoretical formulations of photonic crystals and waveguides. I was greatly inspired by his work, which forms the basis for many of the theoretical results presented in this thesis. I would also like to thank him for the useful discussion and advice.

I am especially grateful to my exceptional MIT professors for their continued support and mentorship, in particular Professors Akintunde Ibitayo Akinwande, Hae-Seung Lee, and Jesus Del Alamo. I would also like to thank Professors Senthil Todadri, Rajeev Ram, Pablo Parrillo and Luis Velasquez, for their great classes that directly contributed to my research.

I am truly indebted to Dr. Zoran Krivokapic (Globalfoundries) for his help and collaboration on the FinFET RBT project. Additionally, I thank Globalfoundries for funding the fabrication of these devices.

Moreover, I would like to thank my colleagues in the Computational Prototyping Group: Zheng Zhang, Zohaib Mahmoud, Lily Weng and Yu-Chung Hsiao. They always inspired me with their discussions and great support.

My work at MIT would not have been possible without the help of MTL support and administrative staff. I would like to thank Vicky Diadiuk, Paul Tierney, Bob Bicchieri, Dave Terry and Paul McGrath for their help with manufacturing processes and fab training. Special thanks goes to Michael McIlrath and Bill Maloney for their continued CAD and IT support. I would also like to thank Valerie DiNardo and Steven O'Hearn for their help and support.

I would like to thank all my collaborators from the DARPA UPSIDE program; in particular: Dr. Narayan Srinivasa, Dr. David Payton, Dr. Matthew Pufall and Prof. Steve Levitan. I would also like to thank my collaborators from the NSF NEEDS project, Professors Mark Lundstrom and Xufeng Wang. I am particularly indebted to Prof. Jaijeet Roychowdhury and Tianshi Wang for their remarkable efforts, support and help with compact models development.

I am also grateful to my professors who introduced me to the integrated circuits, analog design and MEMS devices. In particular, I would like to thank Professors Hani F. Ragai, Mourad El-Gamal, and Emad Hegazi.

Finally, my cordial thanks goes to my family for all their support throughout my life. I would like to thank my parents and sisters for their unconditional love and support throughout the years. Also, I am extremely grateful to my wife Jacqueline Phelesteen for all her love, encouragement, assistance and sacrifices. My achievements would not have been possible without her. I highly appreciate her going the extra mile for the sake of following my dreams.

The revolutionary results of this work were possible through the generous funding of the DARPA UPSIDE program and NSF NEEDS program.

THIS PAGE INTENTIONALLY LEFT BLANK

Contents

THIS PAGE INTENTIONALLY LEFT BLANK

List of Figures

THIS PAGE INTENTIONALLY LEFT BLANK

List of Tables

Chapter 1

Introduction

1.1 Motivation

Digital electronics are becoming increasingly ubiquitous in human civilization. Their unique ability to provide tremendous computational capabilities into small form factor with low power consumption has brought unprecedented versatility, precision and speed to uncountable applications in our world. Furthermore, the exceptional advancement in wireless radio frequency (RF) communications over the past 2 decades enabled the interconnection of diverse devices and systems in previously unforeseen ways. Computers, cell phones, media devices, automobiles, airplanes, navigation systems, medical equipment, home appliances, radars, and industrial machinery are a few classic examples of applications completely revolutionized by the continuous advancement of RF and digital electronics. The interconnectedness of these devices and the continuous access to live information and data have made these devices smarter and more efficient.

Scaling and integration of electronics and communication systems are the keys to their unparalleled advantages. The ability to scale down the size of electronic circuits and integrate twice as many devices each year on a single silicon chip resulted in faster and more efficient circuits and systems that shape the way we live today. Such scaling was predicted by Intel's Gordon Moore more than 5 decades ago. Today's smart phones house far more computational power than Apollo's computers that put a human on the moon. They can also transfer data wirelessly orders of magnitude faster with less errors and provide their users with Internet access anytime on the go.

The tremendous scaling of electronics and RF communication systems (in terms of size, higher operating frequency and power efficiency) has enabled new applications that continue to drive the demand for even more aggressive scaling and integration. Smart wireless sensor networks, wearable electronics, chronic medical implants, smart prosthetics and autonomous robots for civil and military applications are few examples. These systems are battery operated, with stringent battery life requirements; hence the demand for low power electronics. The nature of these applications also calls for small size and light weight, which encourages further integration and scaling. Furthermore, new radio paradigms have started to emerge such as software-defined radios as well agile and cognitive radios. This adds the extra requirements of hardware reconfigurability and agility.

Both digital and RF communication circuits share the requirement of a good frequency source, or oscillator. Oscillators are required to provide a *clock* for digital circuits and act a local oscillator in RF communication systems. High quality factor Q sharp filters are an indispensable component of high purity oscillators. Higher Q is directly reflected as lower noise in the oscillator and/or power consumption. Filters are also used as standalone components in RF systems to select different frequency bands and channels. The ability to integrate high-Q filters inside integrated circuits (ICs) in a compact form factor and low power solution should potentially benefit every device that relies on digital circuits or RF data links.

Filters based on mechanical resonance are known to be far superior to electrical filters in terms of their quality factors. Piezoelectric quartz crystals, with $Q > 100,000$ and reasonable frequency stability, are a good example of mechanical filters that dominated oscillators for decades. Miniaturization and integration of quartz crystals in standard IC technology, specifically CMOS, is recognized to be difficult [1, 2]. Furthermore, quartz crystals are typically bulky, requiring extra space on a system board along with dedicated pins for the integrated circuit. They are also limited to MHz-frequencies ($f_\circ < 300\,\mathrm{MHz}$), which makes a phase-locked-loop (PLL) unavoidable to generate GHz frequency signals which in turn complicates the system and increase its power consumption.

On-chip L-C tank circuits have been the common choice for scaling to GHz-frequencies. However, the resulting quality factor is very limited ($Q < 30$) and their resonance frequency is poorly controlled due to the substantial CMOS process variations. They also tend to consume large on-chip prime die area ($\sim 100 \times 100\,\mu\text{m}^2$), increasing the size and cost of the CMOS die and subsequently the overall system. More versatile devices with monolithic IC integration capabilities are required to meet today's demands for frequency scaling, size and power consumption.

1.2 MEMS Resonators

Micro-electro-mechanical systems (MEMS) and resonators are potential candidates to satisfy today's technological demands [1, 3–5]. MEMS resonators span a wide frequency range from 100 kHz to over 10 GHz. They achieve quality factors exceeding 10^4, within a footprint that is $10^3\times$ smaller than that of on-chip L-C tank circuits. Finally, they have the potential for monolithic integration with commercial CMOS IC technologies, and many research efforts have explored different possibilities for such integration [6].

1.2.1 Classes of MEMS Resonators

The first and by far the simplest MEMS resonator was the *resonant gate transistor* (RGT) developed by Nathanson et al. in 1967. The RGT structure is shown in Figure 1-1 and consists of a clamped cantilever beam, creating a mechanical resonance structure. The cantilever beam also doubles as the gate of a field effect transistor (FET) that senses the vibration by modulation of its drain current. This first demonstration was followed by a wide class of *flexural* beam resonators. Many classical demonstrations have followed including free-free beams, clamped-clamped beams, and the classical comb-drive flexural resonator [3, 8–11].

Another class of MEMS resonators is the *bulk acoustic wave* (BAW) resonator.

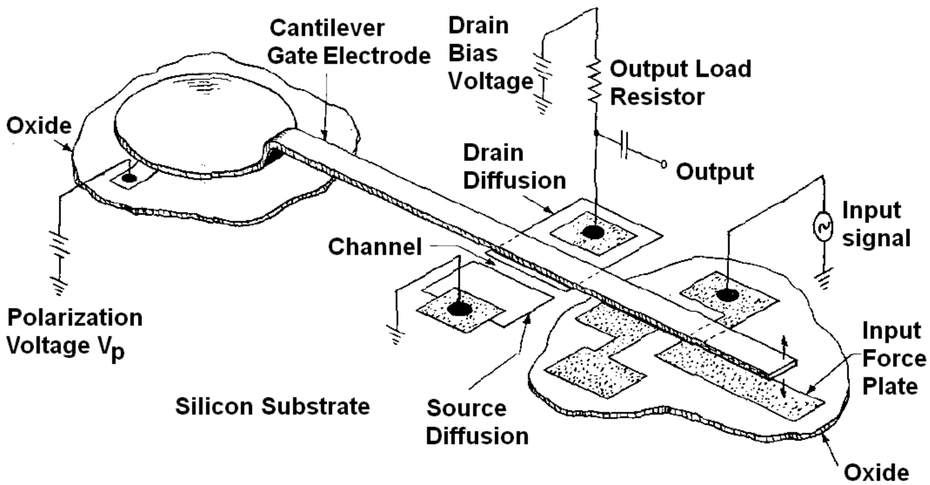


Figure 1-1: The resonant gate transistor developed by Nathason et al. [7] consisting of a gold beam $100\,\mu$m in length, $5-10\,\mu$m thick, and resonating at 5 kHz with a quality factor of 500. (figure after [1]).

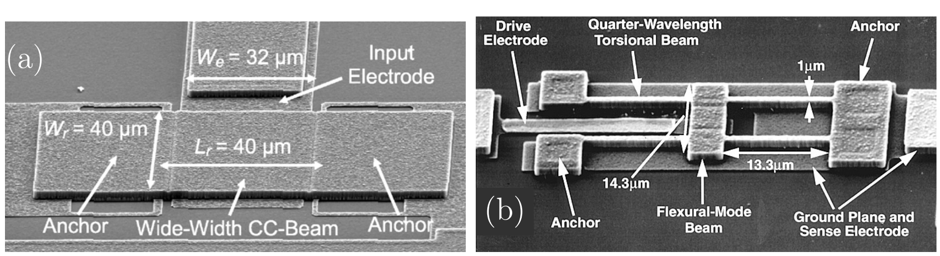


Figure 1-2: Flexural MEMS resonators, (a) clamped-clamped beam after [10] and (b) free-free beam after [9].

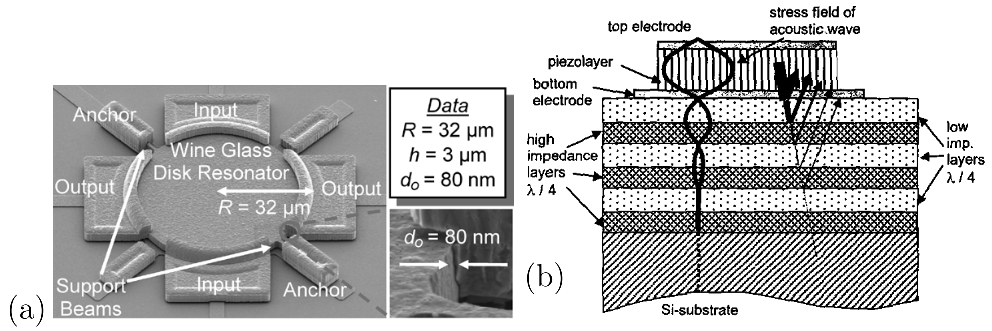


Figure 1-3: Examples of MEMS BAW resonators showing: (a) disk resonator after [10] and (b) a solidly mounted BAW resonator after [16].

These resonators rely on standing shear and longitudinal waves in their operation. They are usually associated with higher frequency due to the higher structural stiffness in these vibrational modes. They are also associated with higher Q compared to flexural modes due to the smaller surface to volume ratio [1]. Disk resonators, Lamè-mode square resonators, Lamb-mode resonators, thin film bulk-acoustic (FBAR) resonator and solidly mounted resonators are all examples of BAW resonators [10,12–16]. Classical examples are shown in Figure 1-3.

1.2.2 Challenges of MEMS Resonators

While providing high Q, small footprint, and potential for integration with CMOS, MEMS resonators carry their own set of challenges.

First, most traditional MEMS resonators involve a *release step*. That is, sacrificial layers are always included during the micro-machining of MEMS resonators. These sacrificial layers get etched away to create the freely suspended and vibrating structure required for resonators and inertial sensors [17]. The release operation is usually a critical step in the MEMS fabrication flow and dramatically affects the process yield. This immediately suggests that solidly mounted BAW resonators [15, 16] are favored in terms of process simplicity.

Next, as mechanical devices, MEMS resonators involve *free surfaces* that vibrate and move; which makes many MEMS resonators sensitive to ambient pressure and humidity adsorption, as well as particle deposition. This directly affects the quality factor and the resonance frequency of these resonators. Also, some resonators rely on *air gaps* for electrostatic actuation. MEMS devices need to be hermetically sealed and protected from the environment to avoid degradation of the device performance. Specialized and costly hermetic and vacuum packaging is required [18–20]. This adds to the overall system size, complexity and cost.

Finally, *true* monolithic integration of MEMS resonators with CMOS circuits have been a challenge so far, as will be discussed in §1.3.1. Integration techniques usually require extensive post-processing or complicated protection at different stages of the fabrication process. Process thermal budget and yield optimization are major issues for these processes.

1.3 MEMS Integration in CMOS

Integration of micro-electro-mechanical (MEM) resonators with CMOS has been studied extensively and several approaches have been proposed, with the most notable including MEMS-first, MEMS-last and CMOS-MEMS technologies [6, 21, 22]. This section describes common integration techniques and the early development that led to the unique integration scheme presented in this thesis.

1.3.1 Integration Strategies

MEMS-First CMOS Integration

MEMS-first integration strategy relies on fabricating the MEMS device before the CMOS electronics [6]. Early realization relied on back-etching the silicon wafer to create membranes for pressure sensors; however, the technique was area inefficient and was discontinued [23]. Techniques relying on deep-reactive ion etching (DRIE) to fabricate devices from silicon-on-insulator (SOI) wafers were also developed, which is the case with Analog Devices' accelerometers [6].

Another technique starts by forming a trench in the wafer [24], when thin-film polysilicon MEMS devices are needed. The poly-silicon structure is then formed with sacrificial layers, and the tench is filled with oxide. The wafer is planarized with chemical mechanical polishing (CMP), and the trench is protected by a nitride passivation film. CMOS electronics are fabricated and are in turn protected. Finally, the nitride sealing is removed and the MEMS structure is released. Alternatively, a process may go as far as releasing the MEMS devices before CMOS fabrication and encapsulating the MEMS in a microshell cavity that also acts as a vacuum packaging for the MEMS devices [25].

While many research groups have demonstrated MEMS-first integration with CMOS, they remain challenging in terms of significantly increased process complexity, protection at intermediate stages, and yield [6].

MEMS-Last CMOS Integration

This approach relies on the fabrication of the MEMS devices after the completion of the CMOS back-end-of-line (BEOL) processing. This technique has the advantage of treating the MEMS process as an optional process module that can be included if desired [6]. This technique is exceptionally useful when the MEMS device is meant to be arrayed with required connection to individual elements as in the case of digital light processing (DLP) technology by Texas Instruments [26] or arrays of capacitive micromachined ultrasonic transducers (CMUTs) for ultrasonic imaging [27]. A notable MEMS above-IC process has been developed by Nabki et al. [28], in which a complete PLL with reference based on a MEMS resonator has been demonstrated [29].

MEMS-last integration still has the advantage over wire bonding solutions of reducing the parasitics, which is highly advantageous especially for MEMS resonators. However, it significantly limits the thermal budget available to the MEMS process in order to protect the CMOS BEOL layers.

CMOS BEOL MEMS Integration

A common class of MEMS devices is implemented by post-processing the CMOS BEOL layers. The work of S.S. Li's group in TSMC CMOS processes is a great example [22, 30, 31], with many MEMS resonators demonstrated in this process. The MEMS structure is actually defined by the metal patterns during BEOL fabrication, just like regular routing for other parts of the CMOS chip. MEMS post-processing starts after the CMOS BEOL processing is completed. The inter-metal dielectric is selectively etched away, creating a released MEMS structure, with the non-MEMS device areas having been passivated earlier for protection. The metal routing in the MEMS structure stops the etching and doubles as electrodes for the MEMS device.

Resonators demonstrated in this kind of process have relied on electrostatic transduction with large air gaps. The latter requires large operating voltages (usually 10's of volts), and the structure still requires challenging packaging for protection from environmental conditions. The demonstrated resonators have been limited to below 100 MHz.

The MEMS resonators considered thus far require complicated and extensive processing, especially when targeting monolithic integration with CMOS. Free moving surfaces, packaging and yield are still major challenges. Creating a high-performance truly solid-state resonator inside commercial CMOS processes is the ultimate goal of this thesis.

1.3.2 Internal Dielectric Transduction

When considering monolithic integration with CMOS, electrostatic transduction becomes appealing due to the absence of piezoelectric materials from the majority of CMOS processes (with ferroelectric RAM—FeRAM—processes being an exception). The air gaps inherent to traditional electrostatic transduction are the major drawback in this case. Air gaps need to be small to achieve better transduction efficiencies at lower voltages, yet small air gaps present a higher risk of stiction during fabrication (Casimir forces makes the effect even more dramatic) [17]. They also become a concern for device hermetic sealing and packaging.

A good alternative for air gap electrostatic transduction is the internal dielectric transduction, first demonstrated by Weinstein and Bhave [32]. By replacing the air gap with a thin dielectric, transduction efficiency increases dramatically due to the thin layer thickness as well as the higher dielectric constant involved. The major contrast in designing internal dielectric transducers is that the dielectric has to be

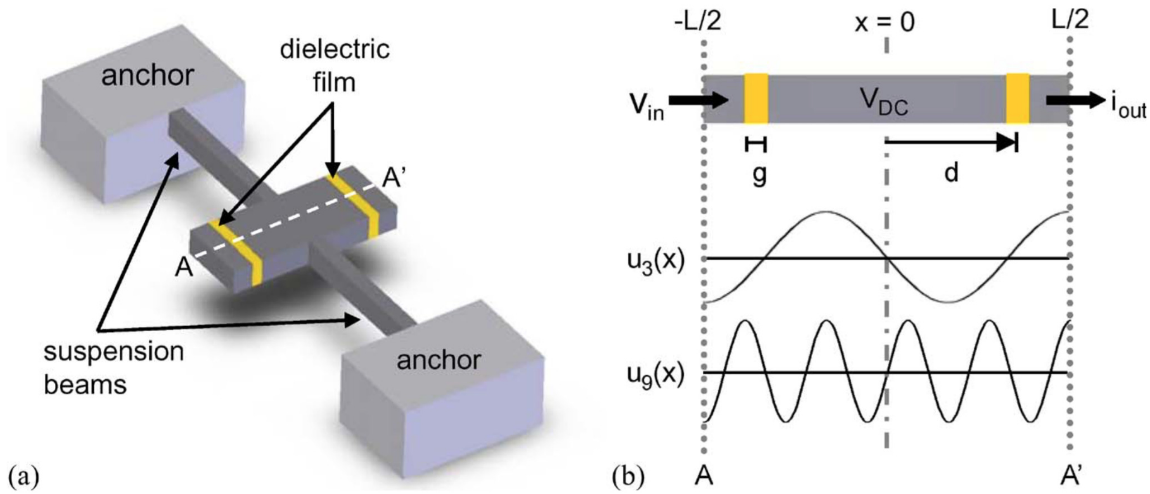


Figure 1-4: Figure after [32] showing (a) MEMS bar resonator based on internal dielectric transduction and (b) displacement showing the mode shape with the dielectric transducer located at the maximum stress points along the bar.

positioned at the location of largest stresses, in contrast to air gap transduction, in which the transducer needs to be at the largest displacement point.

The first MEMS resonator employing internal dielectric transduction is shown in Figure 1-4. This resonator eliminated the air gaps required by electrostatic transduction, making it amenable to CMOS processes.

1.3.3 Acoustic Bragg Reflectors

While internal dielectric transduction solves the problem of air gaps, the resulting resonator still involves free vibrating surfaces. The free surfaces are in fact essential in determining the resonance frequency of the resonator: they set the required boundary conditions to define a given resonance mode. The boundary conditions are only required to be highly reflective in order to create standing elastic waves in a confined volume. The ability to emulate such reflective boundary conditions without free surfaces is the key to creating a completely solid-state MEMS resonator, without any air gaps or exposed moving surfaces. Such a resonator will be immune to environmental conditions and will significantly relax the packaging requirements.

Creating reflective boundary conditions in solid materials for elastic waves can be created by emulating the Bragg reflectors in optics, creating acoustic Bragg reflectors

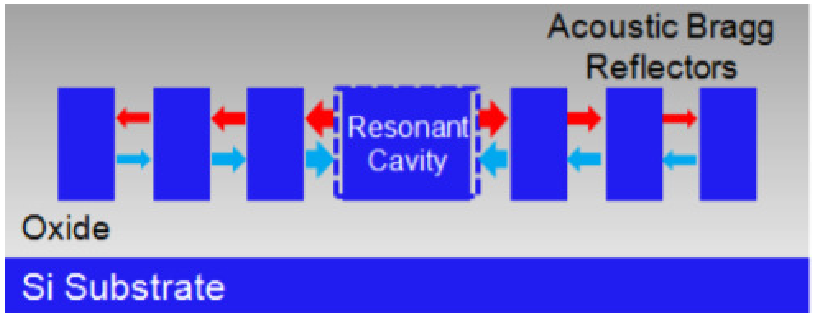


Figure 1-5: Hypothetical MEMS resonator in CMOS technology, using Si/SiO_2 to define a resonance cavity in continuous solid domain after [34].

(ABRs). This idea was first introduced by Newell in 1965 [33], later adopted by Aigner in 2002 [15] and carried to completely unreleased resonators by Wang in 2011 [34]. As in photonics, where individually transparent layers of dielectric can be stacked ($\lambda/4$ thickness each) to create highly reflective boundary conditions, linear elastic materials can be used in a similar setup. Although each material individually supports propagation of elastic waves, the stack with alternating materials of $\lambda/4$ thickness each is highly reflective.

This enables the creation of a truly solid-state MEMS resonator, in which mechanical vibrations can be confined into continuously solid domains without the need for any free surfaces or air gaps. A hypothetical MEMS resonator in CMOS technology is shown in Figure 1-5, where the main resonance cavity is assumed to be silicon, with ABRs on both sides formed from Si/SiO_2 layers [34]. The first realization of solid-state resonators based on these ABRs is discussed in the following section.

1.4 Unreleased CMOS Resonant Body Transistors

Owing to internal dielectric transduction and ABRs, a new class of CMOS-MEMS resonators has emerged. These resonators are truly solid-state without air gaps or free moving surfaces, while operating at GHz-frequencies. They also represent the ultimate integration with CMOS as the resonator is fabricated from the same CMOS layers, just like any other electronic circuit in their. No post-processing or packaging is needed, hence the term *unreleased* MEMS resonators. These devices also incorporate

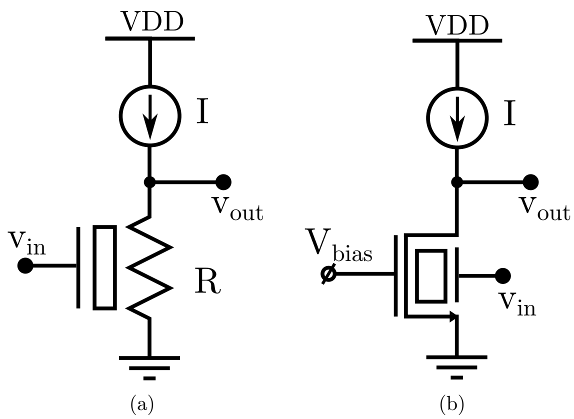


Figure 1-6: Output voltage for (a) piezoresistive sensing, (b) FET sensing and (c) Wheatstone bridge.

active field effect transistor (FET) sensing and hence are referred to as resonant body transistors (RBTs).

1.4.1 Active FET Sensing and Noise Analysis

When considering monolithically integrated MEMS resonators in CMOS, FET sensing becomes a natural choice to harness the high-quality, high-yield MOSFETs available in CMOS technologies. Active transistor sensing for MEMS resonators has been widely demonstrated over different technology platforms, including FETs [35–38] and HEMT sensing [39–42]. Elastic wave stresses modulate the carrier mobility in the transistor channel by virtue of piezoresistivity. Some implementations also involve carrier density modulation when the base transistor material is piezoelectric, or when large strains are induced in the gate dielectric. Both effects create a small signal AC current in the read-out circuit when the FET is biased properly.

Piezoresistive active FET sensing for GHz-frequency resonators provides superior noise performance when compared to using a passive piezo-resistor. First, consider simple piezo-resistor sensing of a MEMS resonator. The sensing resistance variation is proportional to the strain and hence to the resonator input voltage at resonance, as given by

$$\frac{\Delta R}{R} = g_R\, v_{in}, \tag{1.1}$$

where $g_R \in \mathbb{C}$ is a constant. The simplest way to extract the signal from this transducer is to bias it with a constant current and measure the output voltage as depicted in Figure 1-6.a. The output voltage v_{out} and voltage gain A_v are given by

$$v_{out} = g_R\, IR\, v_{in} \quad \text{and} \quad A_v = g_R\, IR, \tag{1.2}$$

where I is the bias current. The output voltage noise power spectral density (PSD) as well as the input referred noise PSD are given by

$$\overline{v_{on}^2} = 4kTR \quad \text{and} \quad \overline{v_{in}^2} = \frac{\overline{v_{on}^2}}{|A_v|^2} = \frac{4kT}{I^2\, g_R^2\, R}. \tag{1.3}$$

Thus, large resistance and large bias current are required to minimize the input referred noise voltage. However, in GHz resonators, the area available to implement the sensing piezo-resistor is usually very limited. As a direct consequence, R is usually limited to small values ($\sim 100\,\Omega$ in the resonators under consideration).

Next, consider active FET sensing. The change in the output current is proportional to the change in the mobility, which is proportional to the strain and hence the input voltage at resonance. A linear approximation for this relation is given by

$$\frac{\Delta I}{I} = g_\mu v_{in}, \tag{1.4}$$

where $g_\mu \in \mathbb{C}$ is a constant. The output voltage and gain from the FET sensing scheme of Figure 1-6.b are given by

$$v_{out} = g_\mu\, I r_\circ\, v_{in} \quad \text{and} \quad A_v = \frac{v_{out}}{v_{in}} = g_\mu\, I r_\circ, \tag{1.5}$$

where $r_\circ$ is the output resistance of the FET. The output thermal noise current PSD as well as the resonator input referred noise are given by

$$\overline{i_{dn}^2} = 4kT\gamma g_m \quad \text{and} \quad \overline{v_{in}^2} = \frac{\overline{i_{dn}^2} r_\circ^2}{|A_v|^2} = \frac{4kT\gamma g_m}{I^2 g_\mu^2}, \tag{1.6}$$

where the MOSFET is assumed to be in saturation and γ is close to unity for short channel devices.

For the same resonator, $|g_\mu|$ and $|g_R|$ are comparable. Using (1.3) and 1.6, for the same bias current, the ratio between the input referred noise is given by

$$\frac{\left.\overline{v_{in}^2}\right|_{FET}}{\left.\overline{v_{in}^2}\right|_{Piezo}} = \gamma g_m R << 1. \tag{1.7}$$

In typical resonators considered here, R is going to be much smaller than $1/g_m$. Thus, FET sensing usually outperforms passive piezo-resistor sensing in terms of noise.

Furthermore, active FET transducers are not limited by the available area. In fact, MOSFET gates set the minimum feature size of the technology, and FET sensing works perfectly even in very small resonators. FET sensing will be considered henceforth for GHz-frequency sensing in CMOS resonators, making them *CMOS-RBTs*.

1.4.2 The First Unreleased CMOS-MEMS RBT

The first implementation of a truly solid-state *unreleased* RBT in IBM 32 nm SOI CMOS technology is reported in [36]. The device is shown in Figure 1-7, along with its measured frequency response. ABRs were used on both sides of a silicon active area slab to create the resonance cavity. Internal dielectric transduction was used for actuation and piezoresistive active FET sensing for signal detection. This RBT showed a quality factor of 30 at 11.1 GHz, with many spurious modes.

The low-Q is mostly attributed to *radiation* losses. The ABRs provides reflection only in the resonator plane, within a small solid angle, whereas elastic waves are allowed to radiate from the cavity from the top and bottom. Furthermore, the strict DRC requirements of the FEOL layers forced the first set of ABRs to be placed at $3\lambda/4$ as opposed to $\lambda/4$, which further reduced the reflection solid angle. In addition, the internal dielectric driving transducer spanned half the cavity width, resulting in major scattering due to the non-uniformity along this direction. Finally, the random metal patterns generated by the CMP fill is believed to be responsible for all the spurious resonances observed in the measured spectrum.

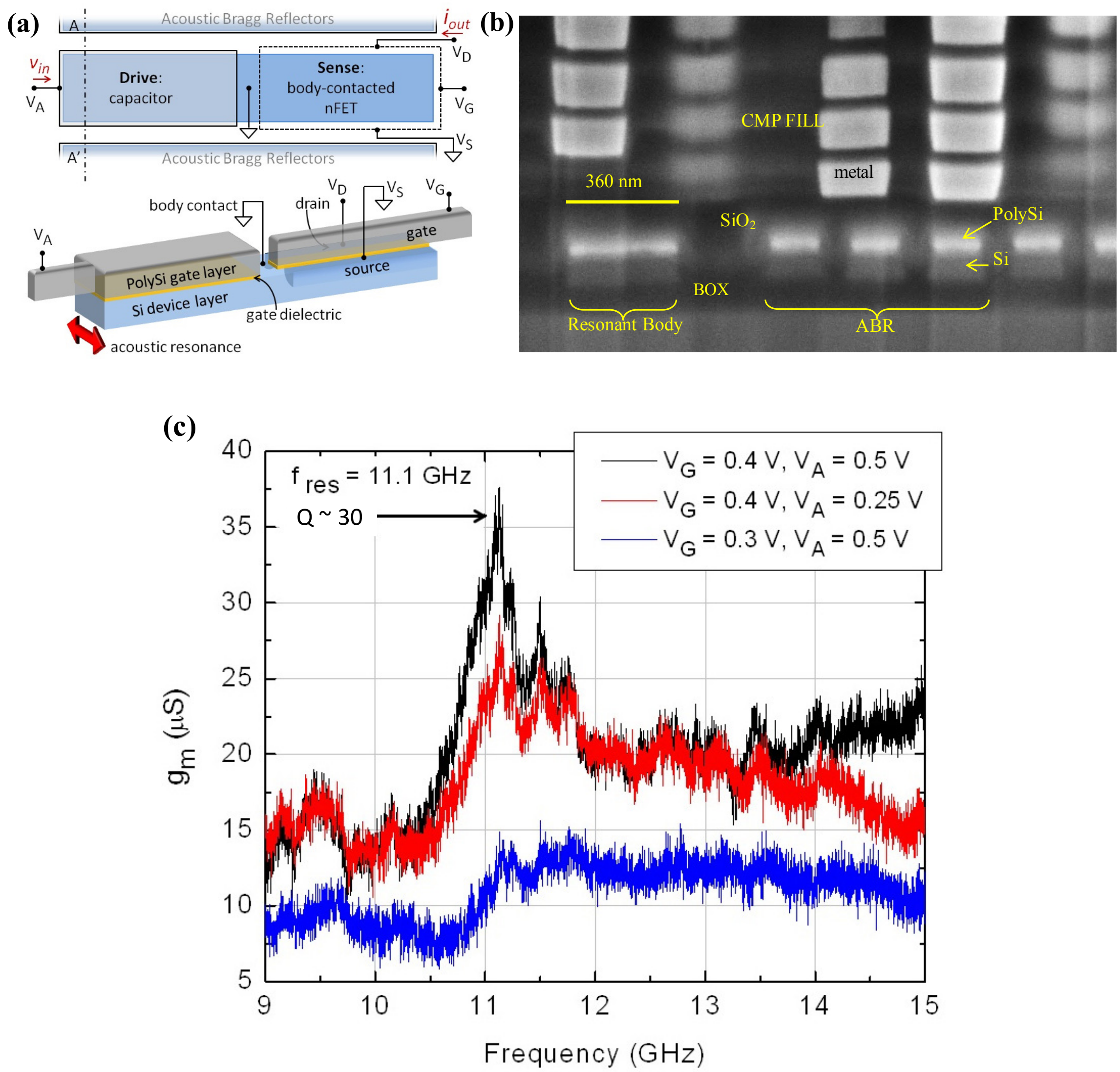


Figure 1-7: First unreleased CMOS MEMS resonator in IBM 32 nm SOI technology using ABRs from the gate stack and FEOL dielectrics. Figure after [36] showing (a) top view of the resonator, (b) SEM cross-section and (c) measured electromechanical transconductance.

In this thesis, superior confinement techniques relying on phononic crystals and phononic waveguides will be presented, implemented and tested. These techniques enable complete energy confinement in all directions, dramatically increasing the quality factor of the resonator and reducing spurious modes.

1.5 GaN-MMIC Technology

Gallium nitride (GaN) is becoming an increasingly popular material for RF monolithic microwave ICs (MMICs) and power electronics. With a Wurtzite crystal structure, GaN is a piezoelectric material with favorable electromechanical properties [43]. The 2D electron gas (2DEG) characteristic to the GaN/AlGaN heterostructure provides high sheet carrier concentration for high electron mobility transistors (HEMTs) and also allows for unique transduction capabilities [44]. With the ever-increasing demand for systems with higher efficiency and smaller footprints, monolithic GaN high-Q filters and low phase-noise oscillators are highly desirable.

Multiple groups have demonstrated MEMS resonators in GaN MMIC technology, where the 2DEG of GaN HEMTs was even used as transducers for driving and sensing acoustic vibrations [13,44–51]. A gold-free (Au-free) GaN MMIC-MEMS process was developed at MIT Microsystems Technology Laboratories (MTL) by L.C. Popa [44]. High-Q Lamb-mode resonators, fabricated side-by-side with HEMTs, were demonstrated in this technology.

In this thesis, the first monolithic MEMS-based oscillator in GaN MMICs is demonstrated, with fabrication and testing performed at MTL. The implemented oscillators show exceptional performance in terms of phase noise and power consumption [52]. Characterization and modeling of the in-house HEMTs were among the challenges of such implementation. Design and characterization of these oscillators are thoroughly discussed in chapter 6.

1.6 Thesis Outline

The goal of this thesis is to present high-Q MEMS resonators, oscillators and filters in standard IC technologies. Unreleased high-Q monolithic CMOS RBTs with no post-processing or special packaging requirements are first demonstrated. The presented RBTs are to be fabricated side-by-side with circuits in commercial CMOS processes without any modification to the process. Novel phononic waveguides in CMOS will be used to achieve complete mechanical energy confinement in the solid CMOS die. The thesis also presents the first monolithic MEMS-based oscillator, as well as lattice and ladder filters in standard GaN MMIC technology. All presented implementations are serious steps towards low-cost and low-power on-chip high-purity oscillators and high-Q filters for ultra-compact digital and RF systems.

The thesis starts in chapter 2 with an operator-theoretic treatment of phononic crystals (PnCs), which are periodic structures that can strongly reflect acoustic waves. The first PnC implementation in multiple CMOS technologies is demonstrated. The effect of the different process variations on PnCs performance is also studied. The chapter explores the potential use of these structures to achieve better confinement in CMOS RBTs.

Next, in chapter 3, PnCs are used to construct phononic waveguides in solid-state CMOS. Practical demonstration is presented for the first RBT that is based on such phononic wave guides. Waveguide perturbation is then studied in an abstract operator-theoretic framework with explicit analogy to photonics and quantum mechanics. The adiabatic theorem is presented with an emphasis on the importance of scattering on the performance of prospective CMOS-RBTs. After developing a thorough understanding of phononic waveguides, the chapter concludes with recommendations on their usage to construct high-Q cavities.

Chapter 4 implements all the concepts developed in the previous chapter to achieve an RBT structure that is capable of achieving good energy confinement and high-Q values. A numerical framework is presented for fast and efficient optimization of the full RBT structure, to maximize quality factor and output signal. Several practical RBT implementation are demonstrated with resonance frequency around 3 GHz and quality factors $Q > 10,000$. Design considerations and testing setup are thoroughly discussed. Also, RBTs in FinFET CMOS technologies are explored, showing resonance up to 33 GHz.

Chapter 5 discusses RBTs compact modeling in Verilog-A, allowing circuit designers to integrate them in their circuits and systems. A modular approach is presented to model the tightly coupled physics of the RBT while allowing flexibility to expand and modify the model as needed.

Chapter 6 explores the integration of MEMS resonators in standard GaN-MMIC technologies, based on a process developed by L.C. Popa at MTL. The process enables the co-fabrication of HEMTs side-by-side with MEMS resonators on the same die. The in-house fabricated HEMTs are characterized and modeled using the MIT virtual source (MVS) model. Next, resonators and HEMTs are used to demonstrate the first monolithically integrated MEMS-based oscillator in GaN MMIC technology. Lattice and ladder filters are also presented, enabling full monolithic RF front-ends in GaN MMICs.

Chapter 7 comprises the conclusion and future work.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 2

Phononic Crystals in CMOS

Phononic crystals (PnCs) are periodic composite structures with special dispersion characteristics that can be used to control the propagation of elastic waves through them [53–65]. *Bandgaps* are distinctive features of PnCs' dispersion relations that, within a certain frequency band, prohibit the propagation of elastic waves in one or more directions. Through proper selection of materials, geometry and dimensions, PnC bandgaps can be engineered to provide high reflectivity for waves propagating in one or all directions. This makes PnCs very attractive for confining acoustic waves, a critical requirement of high-Q resonators and filters [59–62,66,67]. In this chapter, the theoretical foundations for PnCs are presented. Next, the implementation of PnCs in CMOS back-end-of-line (BEOL) materials (Figure 2-1) is demonstrated. This structure is intended to provide superior vertical acoustic confinement for CMOS-RBT, as will be discussed in chapter 3.

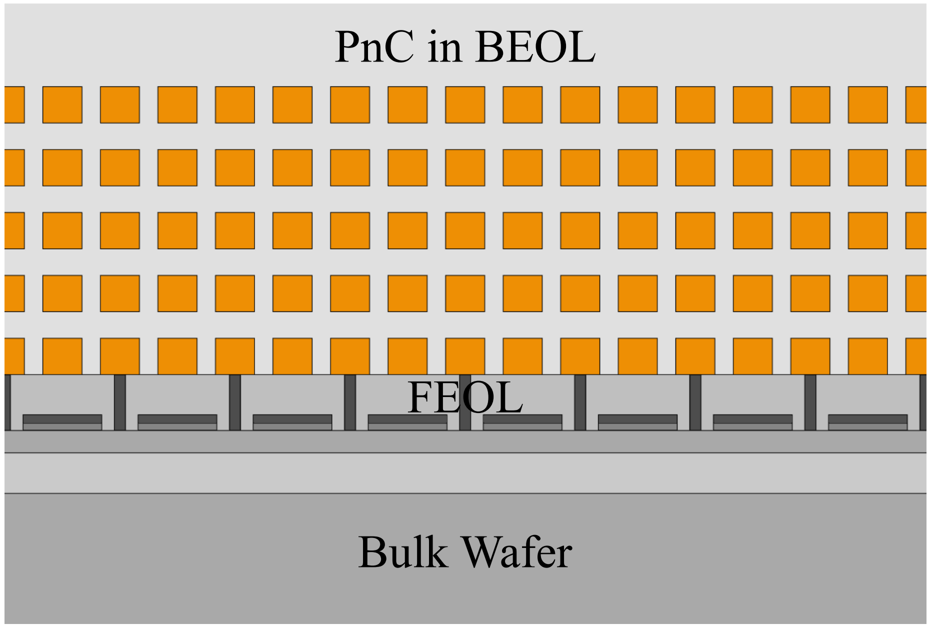


Figure 2-1: Cross-section of a phononic crystal (PnC) implemented as periodic metal stripes in the back-end-of-line (BEOL) of CMOS technology above the front-end-of-line (FEOL) and bulk silicon wafer.

2.1 Phononic Crystals Theory

2.1.1 Linear Elastic Wave Equation

For sufficiently small strains, all the solid materials considered in this work can be accurately studied as linear elastic materials. The constitutive stress-strain relation of such materials is given by Hooke's law [68],

$$T = c : S \quad \Rightarrow \quad T_{ij} = c_{ijkl} S_{kl} \quad , \quad \forall i,j,k,l \in \{1,3\}; \tag{2.1}$$

where T and S are the second rank stress and strain tensors (with 9 elements each), respectively, and $c = c(\vec{r})$ is a fourth rank tensor (with 81 elements) representing the stiffness coefficients of the material. Owing to the symmetry of T and S, c shows multiple symmetries

$$c_{ijkl} = c_{ijlk} = c_{jikl} = c_{jilk}. \tag{2.2}$$

Elastic energy conservation also requires that [68]

$$c_{ijkl} = c_{klij}. \tag{2.3}$$

These symmetries reduce the number of independent parameters in c to 21. In which case using Voigt abbreviated notation is more convenient than full tensor notation [68]. In such notation, T and S in (2.1) are reduced to column vectors, whereas c reduces to a 6×6 matrix [68]

$$T = [T_{xx}\ T_{yy}\ T_{zz}\ T_{yz}\ T_{xz}\ T_{xy}]^{\mathrm{T}}, \tag{2.4a}$$

$$S = [S_{xx}\ S_{yy}\ S_{zz}\ S_{zy}\ S_{xz}\ S_{xy}]^{\mathrm{T}}, \tag{2.4b}$$

$$T_I = C_{IJ} S_J \quad \forall I, J \in \{1, 2, \cdots, 6\}, \tag{2.4c}$$

where the superscript $^{\mathrm{T}}$ indicates transpose operation. The kinematic and dynamic behavior can be described in Voigt notation as [68]

$$\frac{\partial S}{\partial t} = \nabla_s \vec{v} \tag{2.5a}$$

$$\frac{\partial}{\partial t}(\rho \vec{v}) = \nabla \cdot T = \nabla \cdot (c : S) \tag{2.5b}$$

where $\vec{v} = \vec{v}(\vec{r})$ is the velocity vector field (3×1 column vector) and $\rho = \rho(\vec{r})$ is the material density. The divergence $\nabla \cdot$ and symmetric gradient ∇_s are defined as

$$\nabla \cdot = \begin{bmatrix} \partial_x & 0 & 0 & 0 & \partial_z & \partial_y \\ 0 & \partial_y & 0 & \partial_z & 0 & \partial_x \\ 0 & 0 & \partial_z & \partial_y & \partial_x & 0 \end{bmatrix}; \nabla_s = (\nabla \cdot)^{\mathrm{T}} \tag{2.6}$$

For fields with harmonic time dependence of the form $e^{i\omega t}$, the kinematic and dynamic equations (2.5) can be combined to yield the elastic wave equations in solids

$$-\nabla \cdot (c : \nabla_s \vec{u}) = \omega^2 \rho \vec{u} \tag{2.7a}$$

$$-\nabla \cdot (c : \nabla_s \vec{v}) = \omega^2 \rho \vec{v} \tag{2.7b}$$

where $\vec{u} = \vec{u}(\vec{r})$ is the displacement field. Equations (2.7) are linearly dependent as $\vec{v} = i\omega\,\vec{u}$, and either of them provides the full solution to the system.

2.1.2 Operator-Theoretic Formulation for Elastic Wave Equation

In subsequent analysis, the abstract Dirac notation of a *state ket* $|\psi\rangle$ is used to refer to the field solution of (2.7) [69]. Each equation of (2.7) is a generalized Hermitian eigenproblem of the form

$$\hat{A}\,|\psi\rangle = \lambda\hat{B}\,|\psi\rangle \tag{2.8}$$

with $\hat{A} = -\nabla\cdot c : \nabla_s$, $\hat{B} = \rho$ and $\lambda = \omega^2$. $\hat{A}$ and $\hat{B}$ are Hermitian operators under the inner product defined as

$$\langle\vec{\mu}|\vec{\nu}\rangle = \int_\Omega d\Omega\,\vec{\mu}^* \cdot \vec{\nu} \tag{2.9}$$

where * denotes complex conjugation and Ω is the entire solution domain. A proof of the hermiticity of $\hat{A}$ is provided in [70]. The solutions of (2.8) form a complete and orthogonal set of bases for a Hilbert space $\mathcal{H}$ over $\mathbb{C}$ under the inner product $\langle\vec{\mu}|\vec{\nu}\rangle_{\hat{B}}$, with the latter defined as

$$\langle\vec{\mu}|\vec{\nu}\rangle_{\hat{B}} = \int_\Omega d\Omega\,\vec{\mu}^* \cdot \hat{B}\vec{\nu}. \tag{2.10}$$

Considering the inner product of an eigenmode velocity field $|\vec{v}\,\rangle$, we have

$$\langle\vec{v}\,|\vec{v}\,\rangle_{\hat{B}} = \omega^2\,\langle\vec{u}\,|\vec{u}\,\rangle_{\hat{B}} = \int_\Omega d\Omega\,\rho\,\vec{v}^{\,*}\cdot\vec{v}\,. \tag{2.11}$$

This inner product physically corresponds to the kinetic energy* of the eigenmode $|\vec{v}\,\rangle$. Orthogonality implies that two different eigenmodes $|\vec{v}_i\rangle$ and $|\vec{v}_j\rangle$ will have $\langle\vec{v}_i|\vec{v}_j\rangle_{\hat{B}} = 0$ for all $i \neq j$. For non-degenerate eigenmodes, orthogonality physically represents the fact that the energy of each mode is preserved i.e., different modes are not allowed to exchange energy between them. For degenerate eigenmodes, it is

*Formulating (2.7) in terms of $\vec{T}$ yields an inner product that is proportional to the potential energy of the mode.

always possible to find linear combinations thereof that satisfy orthogonality under (2.10) [71].

2.1.3 Symmetry, Translation and Wave Vectors

The physical structure of the problem may often involve a certain symmetry such as mirror, inversion, or translational symmetry. Let such symmetry be associated with an operator $\hat{O}$, such that applying $\hat{O}$ to a solution $|\psi\rangle$ produces a solution satisfying the symmetry criteria. For example, consider an odd eigemode $|\psi\rangle$ with respect to mirroring along a certain axis; then we have $\hat{O}\,|\psi\rangle = -\,|\psi\rangle$.

Applying the symmetry operator $\hat{O}$ before or after the operator $\hat{A}$ must yield the same result; thus, for the structure to be symmetric under $\hat{O}$, the master operator $\hat{A}$ must satisfy [71]

$$\hat{A} = \hat{O}^{-1}\hat{A}\hat{O} \Rightarrow \hat{O}\hat{A} - \hat{A}\hat{O} = 0 \Rightarrow [\hat{O}, \hat{A}] = 0, \tag{2.12}$$

where $[,]$ denotes the commutator bracket. The same condition applies for $\hat{B}$. For an eigenmode $|\psi\rangle$ with eigenvalue λ, owing to the abovementioned commutation relation, one can write

$$\hat{A}(\hat{O}\,|\psi\rangle) = \hat{O}(\hat{A}\,|\psi\rangle) = \hat{O}(\lambda\hat{B}\,|\psi\rangle) = \lambda\hat{B}(\hat{O}\,|\psi\rangle). \tag{2.13}$$

Thus, if $\hat{O}$ is a symmetry operation and $|\psi\rangle$ is an eigenmode corresponding to eigenvalue λ, then $\hat{O}|\psi\rangle$ is also an eigenmode of the system with the same eigenvalue. In the absence of degeneracy, $\hat{O}|\psi\rangle$ must be proportional to $|\psi\rangle$ with some constant $a \in \mathbb{C}$, which implies that $|\psi\rangle$ is also an eigenmode of $\hat{O}$. This is generally true even with degeneracy, as operators that commute share the same set of eigenmodes, and can be classified according to the eigenvalues of both operators [69, 71]. Since symmetry operators are required to commute with $\hat{A}$ and $\hat{B}$ of (2.8), they share the same set of eigenmodes with $\hat{A}$ and $\hat{B}$, and modes can be classified according to their symmetry characteristics. The familiar classification of even and odd modes in symmetric waveguides or quantum wells is a good example.

Symmetry operations are also required to preserve the stored kinetic and strain energy, which are proportional to the inner products $\langle v|v\rangle_{\hat{B}}$ and $\langle S|T\rangle$, respectively [68]. For example, an even mode of a symmetric system cannot change its energy upon mirroring. Formally, this is expressed as

$$\left\langle \hat{O}\psi \middle| \hat{O}\psi \right\rangle = \langle\psi| \hat{O}^{\dagger}\hat{O} |\psi\rangle = \langle\psi|\psi\rangle . \tag{2.14}$$

This requires the symmetry operators to be unitary (or anti-unitary), i.e., $\hat{O}^{\dagger}\hat{O} = I$, and hence $\hat{O}^{\dagger} = \hat{O}^{-1}$, which is the well-known Wigner theorem [72].

Since PnCs are periodic structures, we are primarily concerned with translational symmetry. A translation operator $\hat{T}(\vec{d})$ may be defined as

$$\hat{T}(\vec{d})\, |\vec{u}\,(\vec{r})\rangle = \left|\vec{u}\,(\vec{r} - \vec{d}\,)\right\rangle . \tag{2.15}$$

Since a translation by $\vec{a}$, followed by $\vec{b}$ is equivalent to a translation $\vec{a} + \vec{b}$, $\hat{T}(\vec{d})$ must satisfy (in addition to unitarity) the composition property

$$\hat{T}(\vec{a} + \vec{b}) = \hat{T}(\vec{a})\hat{T}(\vec{b}). \tag{2.16}$$

The translation operator is also required to reduce to the identity as $\vec{d} \to 0$

$$\lim_{\vec{d}\to 0} \hat{T}(\vec{d}) = 1. \tag{2.17}$$

It is easy to show that, to a first order, an infinitesimal translation operator of the form

$$\hat{T}(\Delta\vec{x}) = 1 - i\,\hat{K} \cdot \Delta\vec{x}, \tag{2.18}$$

with $\hat{K}$ being a Hermitian operator, satisfies the unitarity, composition and reduction to identity requirement [69]. For finite translation $\vec{d}$, the translation operator becomes

$$\hat{T}(\vec{d}) = \lim_{N\to\infty} \left(1 - i\,\hat{K} \cdot \frac{\vec{d}}{N}\right)^{N} = \exp\Big(-i\,\hat{K} \cdot \vec{d}\Big). \tag{2.19}$$

The operator $\hat{K}$ is the generator of translation in this formulation, in an explicit analogy to the momentum in quantum mechanics. It is also evident that $[\hat{T}(\vec{d}), \hat{K}] = 0$.

Let the eigenmodes of $\hat{K}$ be $\left|\vec{k}\right\rangle$such that

$$\hat{K}\left|\vec{k}\right\rangle = \vec{k}\left|\vec{k}\right\rangle. \tag{2.20}$$

Applying the translation operator to $\left|\vec{k}\right\rangle$ eigenmode, we have

$$\hat{T}(\vec{d})\left|\vec{k}\right\rangle = \exp\Big(-i\,\hat{K}\cdot\vec{d}\Big)\left|\vec{k}\right\rangle = \exp\Big(-i\,\vec{k}\cdot\vec{d}\Big)\left|\vec{k}\right\rangle. \tag{2.21}$$

Thus, $\left|\vec{k}\right\rangle$ is also an eigenmode of $\hat{T}(\vec{d})$ with an eigenvalue $\exp\Big(-i\,\vec{k}\cdot\vec{d}\Big)$. For plane waves, a translation by $\vec{d}$ will induce a phase shift that is the dot product of the wave vector and $\vec{d}$. In comparison to (2.21), we identify the eigenvalue $\vec{k}$ of the operator $\hat{K}$ as the wave vector and $\left|\vec{k}\right\rangle$ as a plane wave solution. We will refer to $\left|\vec{k}\right\rangle$ as the wave vector basis, and $|\vec{u}\,(\vec{r})\rangle$ as the displacement basis for the solution of the wave equation.

To find the relation between $\left|\vec{k}\right\rangle$ and $|\vec{u}\,(\vec{r})\rangle$, we start by applying an infinitesimal displacement operator to an arbitrary elastic wave $|\psi\rangle$ as follows

$$\begin{aligned}
\hat{T}(\Delta\vec{x})\,|\psi\rangle &= \Big(1 - i\,\hat{K}.\Delta\vec{x}\Big)\,|\psi\rangle \\
&= \int d\vec{u}\quad |\vec{u}\,(\vec{r})\rangle\langle\vec{u}\,(\vec{r})|\;\hat{T}(\Delta\vec{x})\,\Delta\vec{x}\,|\psi\rangle \\
&= \int d\vec{u}\quad |\vec{u}\,(\vec{r})\rangle\;\langle\vec{u}\,(\vec{r}+\Delta\vec{x})|\psi\rangle \\
&= \int d\vec{u}\quad |\vec{u}\,(\vec{r})\rangle\Bigg(\;\langle\vec{u}\,(\vec{r})|\psi\rangle + \Delta\vec{x}\cdot\boldsymbol{\nabla}\,\langle\vec{u}\,(\vec{r})|\psi\rangle\;\Bigg),
\end{aligned} \tag{2.22}$$

where we have used the completeness relation $\int d\vec{u}\quad |\vec{u}\,(\vec{r})\rangle\langle\vec{u}\,(\vec{r})| = 1$ and the unitarity of $\hat{T}(\Delta\vec{x})$. Taking the inner product with $\langle\vec{u}\,'(\vec{r})|$, we get

$$\langle\vec{u}\,'(\vec{r})|\hat{K}|\psi\rangle = i\boldsymbol{\nabla}\,\langle\vec{u}\,'(\vec{r})|\psi\rangle\,, \tag{2.23}$$

which is in explicit analogy to the displacement and momentum relations in quantum mechanics [69]. Considering the special case of $|\psi\rangle = \left|\vec{k}\right\rangle$, we get

$$\left\langle \vec{u}\,'(\vec{r})\middle|\hat{K}\middle|\vec{k}\right\rangle = \vec{k}\left\langle \vec{u}\,'(\vec{r})\middle|\vec{k}\right\rangle = i\boldsymbol{\nabla}\left\langle \vec{u}\,'(\vec{r})\middle|\vec{k}\right\rangle, \tag{2.24}$$

which leads to

$$\left\langle \vec{u}\,'(\vec{r})\middle|\vec{k}\right\rangle = Ne^{-i\vec{k}\cdot\vec{r}}, \tag{2.25}$$

where $N \in \mathbb{C}$ is a normalization constant that can be determined from the mode power relating to $\langle \vec{u}\,(\vec{r})|\vec{u}\,(\vec{r})\rangle$.

A homogeneous medium represents a continuous translational symmetry. Solutions to (2.7) are simply plane waves with a wave vector $\vec{k}$ and a frequency $\omega = c|\vec{k}|$, where c is the speed of sound for the given wave type, being longitudinal (P-wave) or shear (S-wave). Based on the continuous translational symmetry, $\hat{A}$ from (2.8) must commute with $\hat{T}(\vec{d})$ for all $\vec{d}$. Thus, we can use both $\vec{k}$ and ω to label the eigenmodes $\left|\omega, \vec{k}\right\rangle$ of (2.8) in a medium with a translational symmetry. Moreover, according to Noether's theorem, the eigenvalue $\vec{k}$ of the generator of translation $\hat{K}$ is conserved (both in space and time) by virtue of the translational symmetry [69, 73].

2.1.4 PnC Periodicity and the Bloch Theorem

Consider the periodic PnC structure to be generated by translating a *unit cell* along the lattice vector $\vec{R} = l\,\vec{a}_1 + m\,\vec{a}_2 + n\,\vec{a}_3$, where $\vec{a}_i$ are the primitive lattice vectors and (l, m, n) are integers [71]. By virtue of the discrete symmetry, the translation operator $\hat{T}(\vec{R})$ commutes with $\hat{A}$ and when applied to a plane wave solution $\left|\omega, \vec{k}\right\rangle$ yields

$$\hat{T}(\vec{R})\left|\omega, \vec{k}\right\rangle = e^{-i\vec{k}\cdot\vec{R}}\left|\omega, \vec{k}\right\rangle = \left|\omega, \vec{k}\right\rangle \quad , \quad \forall\, \vec{k}\cdot\vec{R} = 2\pi N. \tag{2.26}$$

Thus all solutions with $\vec{k}\cdot\vec{R} = 2\pi N$ with $N \in \mathbb{Z}$ are degenerate. The wave vector $\vec{k}$ can be assumed to obtain its values from a *reciprocal lattice*, with primitive lattice vectors $\vec{b}_i$ defined as

$$\vec{b}_1 = \frac{2\pi\, \vec{a}_2 \times \vec{a}_3}{\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3}\ ,\ \vec{b}_2 = \frac{2\pi\, \vec{a}_1 \times \vec{a}_3}{\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3}\ ,\ \vec{b}_3 = \frac{2\pi\, \vec{a}_1 \times \vec{a}_2}{\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3}\ , \tag{2.27}$$

and reciprocal lattice vectors $\vec{G} = l'\,\vec{b}_1 + m'\,\vec{b}_2 + n'\,\vec{b}_3$, such that $\vec{G} \cdot \vec{R} = 2\pi N$ [71]. This makes the set of solutions corresponding to $\vec{k} + \vec{G}$ a denumerably infinite set of degenerate solutions for all $l', m', n' \in \mathbb{Z}$. By virtue of the discrete symmetry, $\vec{k}$ is still conserved up to arbitrary $\vec{G}$.

Since a linear combination of the degenerate solutions is itself a solution, one can form the linear sum

$$|\psi\rangle = \sum_{\vec{G}} c_G \left|\omega, \vec{k} + \vec{G}\right\rangle . \tag{2.28}$$

In displacement basis, this can be expressed as

$$\begin{aligned}\langle \vec{u}\,(\vec{r})|\psi\rangle &= \sum_{\vec{G}} c_{\vec{G}} N_{\vec{G}} \exp\Big(-i\,(\vec{k} + \vec{G}) \cdot \vec{r}\Big) \\ &= \exp\Big(-i\,\vec{k} \cdot \vec{r}\Big) \sum_{\vec{G}} c_{\vec{G}} N_{\vec{G}} \exp\Big(-i\,\vec{G} \cdot \vec{r}\Big) \\ &= e^{-i\,\vec{k}\cdot\vec{r}}\ \vec{u}_{\vec{k}}(\vec{r}),\end{aligned} \tag{2.29}$$

where $\vec{u}_k(\vec{r})$ is a periodic function on the PnC lattice (as recognized from its Fourier series expansion above) such that $\vec{u}_k(\vec{r}+\vec{R}) = \vec{u}_k(\vec{r})$. This is *Bloch theorem* for waves in periodic structures and forms the basis of PnC theory [71, 74]. The *first Brillouin Zone* refers to the region in the reciprocal lattice that is closest to $\vec{k} = 0$ and contains distinct modes that can't be obtained from one another by applying a $\vec{G}$ shift to $\vec{k}$.

2.1.5 PnC Structural Symmetry

In addition to the inherent periodicity and translational symmetry, the PnC unit cell itself may possess other symmetries. Rotational symmetry (including rotation, mirror and inversion symmetries) can be associated with an operator $\hat{\mathcal{R}}(\hat{n}, \theta)$ that rotates vectors by angle θ around axis $\hat{n}$. A corresponding operator $\hat{O}_{\mathcal{R}}$ to rotate a vector field $\vec{\mu}$ is defined as

$$\hat{O}_{\mathcal{R}}\,\vec{\mu} = \hat{\mathcal{R}}\,\vec{\mu}\left(\hat{\mathcal{R}}^{-1}\,\vec{r}\right). \tag{2.30}$$

By considering a given rotation symmetry for a PnC, the unitary rotation operator $\hat{O}_{\mathcal{R}}$ commutes with $\hat{A}$ and $\hat{B}$ from (2.8), such that

$$\begin{aligned}\hat{A}(\hat{O}_{\mathcal{R}}\,\vec{u}\,(\vec{r})) &= \hat{O}_{\mathcal{R}}\,(\hat{A}\,\vec{u}\,(\vec{r})) = \hat{O}_{\mathcal{R}}\,(\omega^2(\vec{k})\,\hat{B}\,\vec{u}\,(\vec{r}))\\ &= \omega^2(\vec{k})\hat{B}\,(\hat{O}_{\mathcal{R}}\,\vec{u}\,(\vec{r}))\end{aligned} \tag{2.31}$$

which indicates that $\hat{O}_{\mathcal{R}}\vec{u}\,(\vec{r})$ are eigenmodes with the same eigenfrequency $\omega(\vec{k})$ as $\vec{u}\,(\vec{r})$. As for the translation by a lattice vector $\vec{R}$ we have

$$\begin{aligned}\hat{T}(\vec{R})(\hat{O}_{\mathcal{R}}\vec{u}\,(\vec{r})) &= \hat{O}_{\mathcal{R}}(\hat{T}(\hat{\mathcal{R}}^{-1}R)\vec{u}\,(\vec{r}))\\ &= \hat{O}_{\mathcal{R}}(e^{-i\vec{k}\cdot\hat{\mathcal{R}}^{-1}\vec{R})}\vec{u}\,(\vec{r}))\\ &= e^{-i(\hat{\mathcal{R}}\vec{k})\cdot\vec{R}}\hat{O}_{\mathcal{R}}\vec{u}\,(\vec{r})\end{aligned} \tag{2.32}$$

where we have used $\vec{k}\cdot(\hat{\mathcal{R}}^{-1}\vec{R}) = (\hat{\mathcal{R}}\vec{k})\cdot\vec{R}$ owing to the unitarity of $\hat{\mathcal{R}}$. Thus, the eigenmode $\hat{O}_{\mathcal{R}}\vec{u}\,(\vec{r})$ has the same eigenfrequency $\omega(\vec{k})$ as $\vec{u}\,(\vec{r})$, while being associated with the reciprocal lattice vector $\hat{\mathcal{R}}\vec{k}$, such that $\omega^2(\hat{\mathcal{R}}\vec{k}) = \omega^2(\vec{k})$. Thus, the reciprocal lattice displays the same symmetry as the PnC lattice. The same is true for all of the *point group* symmetries (rotation, reflection and inversion) of the lattice [71].

Moreover, in the absence of losses from the structure, and having $\hat{A}$ and $\hat{B}$ Hermitian, conjugation of (2.8) leads to time reversal symmetry. That is, an observer can not distinguish between a wave with a wave vector $\vec{k}$ traveling forward in time and another with wave vector $-\vec{k}$ with time rolling backwards. Thus, we will always have $\omega(\vec{k}) = \omega(-\vec{k})$ for all phononic crystals considered here.

Owing to such symmetries, only a subset of the first Brillouin Zone referred to as the *Irreducible Brillouin Zone* (IBZ) is sufficient to fully generate the dispersion relation on the entire reciprocal lattice. The extrema of the dispersion relation $\omega(\vec{k})$ will usually occur at the edges of the IBZ [71] for all practical purposes. Thus, when characterizing the PnC bandgaps, it is sufficient to consider solutions on the edges of the IBZ.

2.1.6 Coordinates and Material Transformations

It is also important to consider the effects of coordinate and material transformations on the PnC dispersion relation. Consider a coordinate scaling in the form of a linear mapping $\vec{r} \mapsto s\vec{r}$ for some $s \in \mathbb{R}^+$. Using $\vec{r}\,' = s\vec{r}$ and $\nabla = \nabla'/s$, the master equation (2.7) becomes

$$-\nabla' \cdot c'(\vec{r}\,') : \nabla'_s \vec{u}\,(\frac{\vec{r}\,'}{s}) = \left(\frac{\omega}{s}\right)^2 \rho'(\vec{r}\,')\vec{u}\,(\frac{\vec{r}\,'}{s}) \tag{2.33}$$

where $c'(\vec{r}\,') = c(\vec{r}\,'/s)$ and $\rho'(\vec{r}\,') = \rho(\vec{r}\,'/s)$. Thus, uniform coordinate scaling results in an inversely proportional scaling of the frequency $\omega \mapsto \omega/s$. Moreover, scaling of material properties $c \mapsto s^2 c$ or $\rho \mapsto \rho/s^2$ will result in a similar scaling for the frequency $\omega \mapsto \omega/s$. This very simple conclusion has far-reaching and important implications. It allows a designer to engineer the dispersion characteristics $\omega(\vec{k})$ of a PnC by scaling appropriate dimensions. Moreover, it helps build intuition about perturbations in dimensions or material properties which proves to be very useful in characterizing fabrication variations and structural non-idealities.

2.1.7 PnC Dispersion Relation and Bandgaps

Based on the previous discussions, the PnC dispersion relation, or $\omega - \vec{k}$ relations, can be obtained by solving the eigenvalue problem (2.7) for $\vec{u}_{\vec{k}}(\vec{r})$ in a single unit cell with periodic boundary conditions of the form

$$\vec{u}\,(\vec{r} + \vec{R}) = \exp\Big(i\vec{k} \cdot \vec{R}\Big)\,\vec{u}\,(\vec{r}). \tag{2.34}$$

In this formulation, $\vec{k}$ is a parameter and the eigenvalue problem (2.7) has to be solved for different $\vec{k}$ along the boundaries of the IBZ as discussed in §2.1.5.

To understand why bandgaps appear in the dispersion relation of phononic crystals, we first consider a 1D periodic structure consisting of 2 materials with period

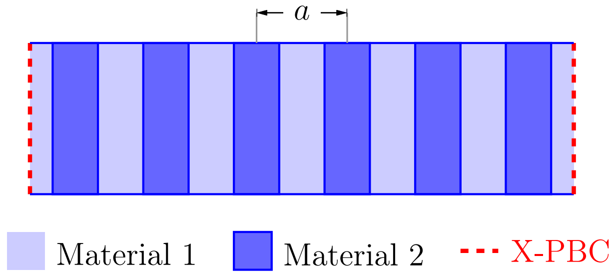


Figure 2-2: 1-D periodic medium with 2 different materials.

a as shown in Figure 2-2 [71]. The width of each material is set to be $a/2$ or half the period. Materials 1 and 2 are assumed to be isotropic with the same density and Poisson's ratio. However, the Young modulus of the materials E_1 and E_2 are allowed to differ. In order to find the dispersion relation $\omega - k_x$ of this structure, the eigenvalue problem (2.7) is solved over a single period with periodic boundary conditions as described in (2.34).

First, we consider the case with $E_1 = E_2$, which translates to a single homogeneous material. The resulting dispersion relation is shown in Figure 2-3.a, which is just a straight line corresponding to $\omega = c\,k_x$, as expected. Only the first Brillouin zone is shown, with clear folding from the neighboring Brillouin zones at $k_x = \pi/a$ and $k_x = -\pi/a$.

Next, consider the case with $E_2 = 0.9\,E_1$; that is, a small mismatch between the two materials is introduced. The dispersion relation for this case is shown in Figure 2-3.b. A discontinuity in the bands at $k_x = \pi/a$ and $k_x = -\pi/a$ starts to appear. Increasing the mismatch to make $E_2 = 0.5\,E_1$, we obtain the dispersion relation in Figure 2-3.c. With a larger mismatch, a wider phononic bandgap is obtained. Moreover, Figure 2-4 shows the normalized T_{xx} stress for the modes in band 1 and band 2 at $k_x = \pi/a$, i.e., the mode just below the bandgap and the one just above it. It is clear that the highest frequency mode in band 1 has higher stress in material 2 with smaller stiffness, whereas band 2 mode has higher stress in material 1 with larger stiffness. To understand why such particular distribution arises, we have to consider the *variational principle* [69, 71].

The variational principle simply states that the mode with the lowest eigenvalue

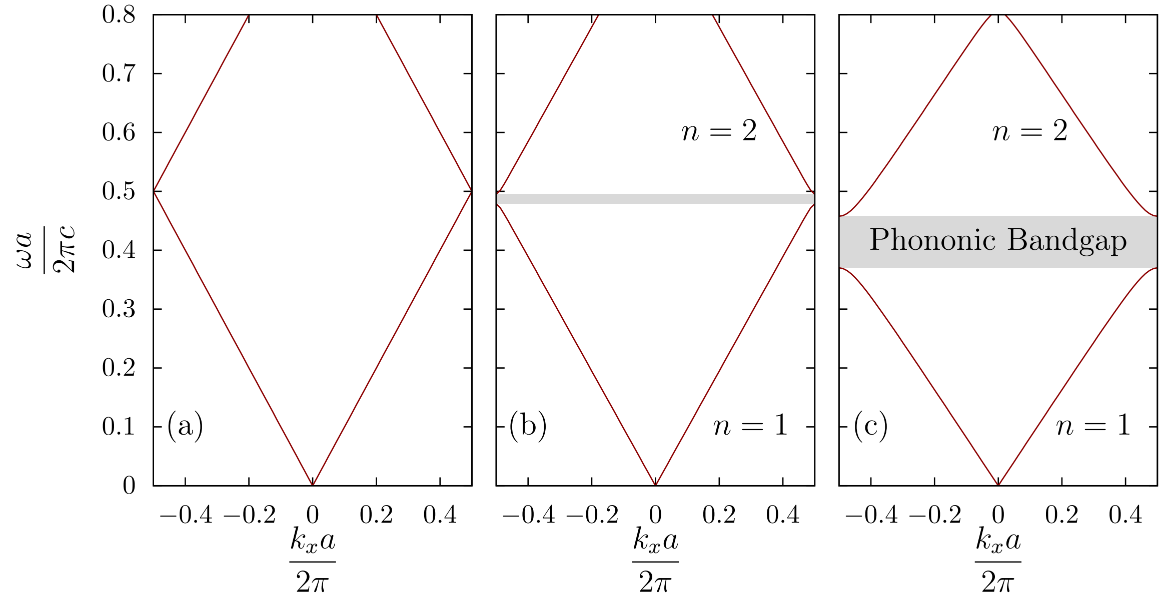


Figure 2-3: Dispersion relation for the periodic structure of Figure 2-2 for (a) $E_1 = E_2$, (b) $E_2 = 0.9\,E_1$ and (c) $E_2 = 0.5\,E_1$.

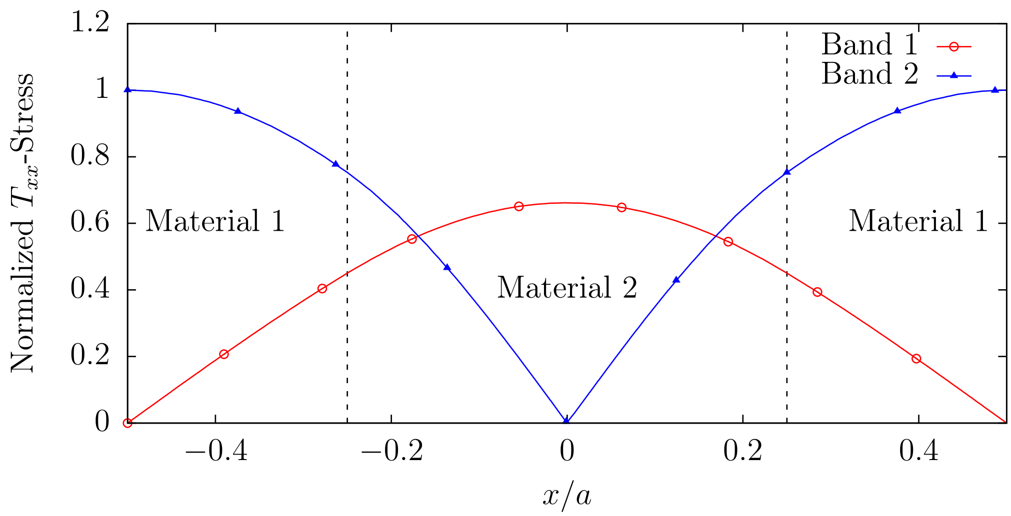


Figure 2-4: Normalized stress $|T_{xx}|$ for the mode just below and just above the bandgap at $k_x = \pi/a$ with material stiffness mismatch set to 0.5.

$\lambda_{\min} = \omega_{\min}^2$ has to minimize the functional $\mathcal{U}(\psi)$ given by

$$\mathcal{U}(\psi) = \frac{\langle\psi|\,\hat{A}\,|\psi\rangle}{\langle\psi|\psi\rangle_{\hat{B}}} \geq \lambda_{\min} \quad \forall\,|\psi\rangle\,, \tag{2.35}$$

known as the Rayleigh quotient [69, 71]. To prove this statement, let the mode with lowest eigenvalue be denoted $|0\rangle$ and consider another mode that has a slightly different field distribution denoted by $\left|\tilde{0}\right\rangle$. The functional $\mathcal{U}(\tilde{0})$ becomes

$$\mathcal{U}(\tilde{0}) = \frac{\left\langle\tilde{0}\right|\hat{A}\left|\tilde{0}\right\rangle}{\left\langle\tilde{0}\middle|\tilde{0}\right\rangle_{\hat{B}}}. \tag{2.36}$$

Expanding $\left|\tilde{0}\right\rangle$ in terms of the eigenmodes $|a\rangle$ of (2.8) with eigenvalues a we get

$$\left|\tilde{0}\right\rangle = \sum_a |a\rangle \left\langle a\middle|\tilde{0}\right\rangle_{\hat{B}}. \tag{2.37}$$

Substituting into 2.36 we get

$$\mathcal{U}(\tilde{0}) = \frac{\sum_a \left\langle\tilde{0}\right| a\hat{B}\,|a\rangle \left\langle a\middle|\tilde{0}\right\rangle_{\hat{B}}}{\sum_a \left\langle\tilde{0}\middle|\hat{B}\middle|a\right\rangle \left\langle a\middle|\tilde{0}\right\rangle_{\hat{B}}} = \frac{\sum_a a\left|\left\langle a\middle|\tilde{0}\right\rangle_{\hat{B}}\right|^2}{\sum_a \left|\left\langle a\middle|\tilde{0}\right\rangle\right|^2}. \tag{2.38}$$

Substituting with $a = a - \lambda_{\min} + \lambda_{\min}$, we get

$$\mathcal{U}(\tilde{0}) = \lambda_{\min} + \frac{\sum_a (a-\lambda_{\min})\left|\left\langle a\middle|\tilde{0}\right\rangle_{\hat{B}}\right|^2}{\sum_a \left|\left\langle a\middle|\tilde{0}\right\rangle\right|^2} \geq \lambda_{\min}, \tag{2.39}$$

where $(a - \lambda_{\min})$ is necessarily positive.

Thus from the variational principle, the mode with lowest eigenvalue and hence lowest eigenfrequency must have field configuration that minimizes $\mathcal{U}(\psi)$. By recalling that $\hat{A} = -\nabla\cdot c : \nabla_s$, it is clear that softer materials with smaller c components will be favored to contain most of the mode strain, and hence most of the mode energy, while satisfying all necessary boundary conditions. Note that $\vec{k}$ is still a parameter and in this case enforces a set of boundary conditions. This discussion applies individually to all values of $\vec{k}$. Next, when considering the following mode, it must minimize

$\mathcal{U}(\psi)$ with the additional constraint of being orthogonal to the previous one and still satisfying all boundary conditions [71]. This is applicable to all subsequent modes, with the constraint that every mode should be orthogonal to all the other modes.

Thus in general, all modes will favor having most of their energy in soft materials to minimize $\mathcal{U}(\psi)$. However, the requirements of being orthogonal to previous modes and satisfying the boundary conditions may enforce a substantially different field distribution, significantly extending into stiffer materials. This results in a step increase in energy, eigenvalue and hence frequency. This behavior is seen in the field configurations of Figure 2-4 and explains the appearance of the bandgap in Figure 2-3.b. It further explains why the bandgap grows with larger mismatch between the material stiffness. More accurately, the bandgap grows with the acoustic impedance ($\sqrt{c/\rho}$) mismatch between the constituting materials. The dimensions and shape of the structure plays a crucial role as well in determining the bandgap as they affect the possible field configurations in the PnC.

2.2 CMOS PnC Implementations

As explained in §2.1.7, the width of the PnC bandgap depends on the acoustic impedance mismatch, shape and dimensions of the unit cell, among other factors [53, 63, 71]. The BEOL layers of CMOS technologies usually have several materials with large mismatch in acoustic impedance that can be leveraged to implement wide-bandgap PnCs [59]. Table 2.1 shows the acoustic impedance for common CMOS BEOL materials. Copper metallization in low-κ SiCOH dielectric background as well as tungsten in SiO_2 are notable candidates for PnC implementation, with impedance contrast ratios on the order of 19$\times$ and 7$\times$, respectively. Such large mismatch, coupled with the small feature size available in CMOS, allows for PnCs with wide bandgaps that are ideal for energy confinement in unreleased resonators.

To guarantee manufacturability and high yield, commercial CMOS technologies impose strict Design Rule Check (DRC) constraints. DRC rules include specifications for metal widths, separations and even restrictions on maximum and minimum filling

Table 2.1: Mechanical properties for popular materials in commercial CMOS technologies.

Material	ρ (kg/m^3)	c_{11} (GPa)	Z_{11} $(MRayl)$	c_{44} (GPa)	Z_{44} $(MRayl)$
Si$< 100 >$	2329	194.3	21.2	79.5	13.6
SiO_2	2200	75.2	12.9	29.9	8.1
SiCOH	1060	3.96	2.05	1.32	1.18
TEOS	2160	49.4	10.3	19.7	6.5
Tungsten	17600	525.5	96.2	160.5	53.1
Copper	8700	176.5	39.1	40.7	18.8
Aluminum	2735	111.1	17.4	28.9	8.9

densities. Moreover, only the horizontal dimensions of the mask layers are available for the designer to control, while all the vertical dimensions are set by the foundry to optimize the fabrication process. This significantly limits the feasible structures and dimensions of a PnC implemented in CMOS.

2.2.1 BEOL PnC in IBM 32 nm SOI

The simplest 2D PnC design in CMOS consists of parallel and periodic rectangular metal lines formed from different BEOL layers. A 3D illustration of this structure in IBM 32 nm SOI technology [75] is shown in Figure 2-5.c. This structure is very favorable from a manufacturability point of view, as it is very similar to parallel bus connections very common in CMOS designs. Moreover, the structure is simple enough to design in full compliance with the DRC rules of the CMOS process.

For this particular PnC, the copper metal stripes have a width and spacing in x-direction of 165 nm and 85 nm, respectively. The corresponding unit cell as well as the IBZ are shown in Figure 2-5.a. The unit cell has two mirror symmetries, namely σ_x and σ_y, which along with time reversal symmetry reduce the first Brillouin Zone to the rectangular IBZ shown in Figure 2-5.b. The IBZ doesn't have the rather familiar triangular shape (associated with square unit cells), due to the lack of the 90° (C_{4v}) rotational symmetry. This is generally the case for most CMOS technologies. The

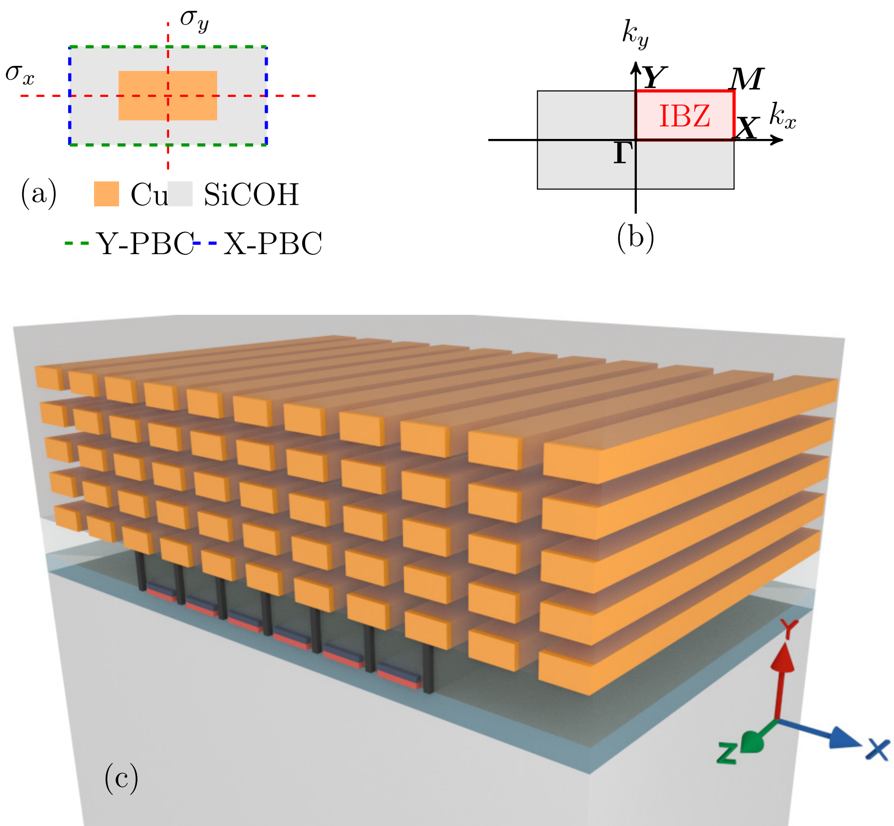


Figure 2-5: (a) Unit cell for a PnC implemented in IBM32SOI technology BEOL, showing the mirror axes σ_x and σ_y. Blue and green dashed lines indicate the periodic boundary conditions applied in FEM simulation. (b) First Brillouin Zone and IBZ of the reciprocal lattice for the unit cell of (a). (c) 3D view of the full PnC on top of an RBT.

unit cell is assumed to have dimensions of a and b in x and y directions, respectively. The primary lattice vectors are given by

$$\vec{a}_1 = a\hat{x} \ , \ \vec{a}_2 = b\hat{y}. \tag{2.40}$$

The corresponding reciprocal lattice vectors are thus given by

$$\vec{b}_1 = \frac{2\pi}{a}\hat{x} \ , \ \vec{b}_2 = \frac{2\pi}{b}\hat{y}. \tag{2.41}$$

As mentioned in §2.1.5, the extrema of the dispersion relation occur at the IBZ edges.

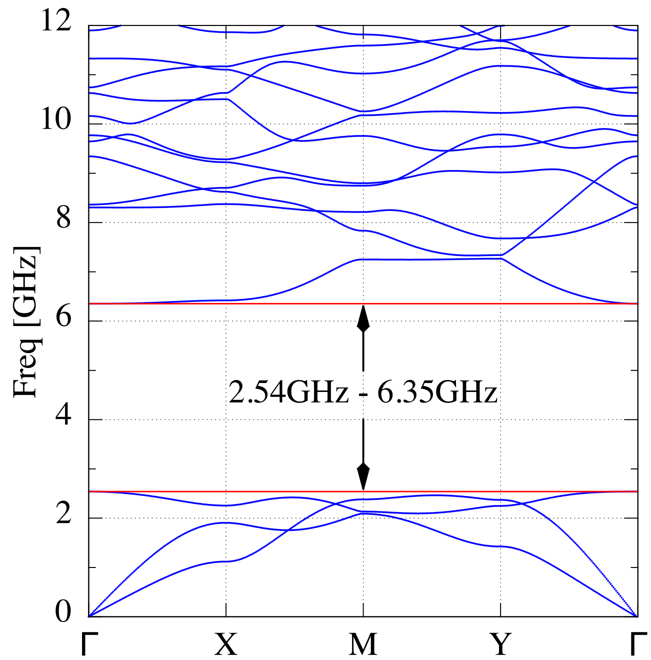


Figure 2-6: Dispersion relation for the PnC of Figure 2-5 implemented in IBM 32 SOI, showing a bandgap of 3.81 GHz (2.54 GHz–6.35 GHz).

It is hence sufficient to consider wave vectors $\vec{k}$ in the sequence

$$\frac{k_x a}{\pi} : 0 \to 1 \to 1 \to 0, \tag{2.42a}$$

$$\frac{k_y b}{\pi} : 0 \to 0 \to 1 \to 0. \tag{2.42b}$$

Dispersion characteristics $\omega - \vec{k}$ of the PnC are found through numerical simulations. Finite element modeling (FEM) is preferred for its accuracy, good convergence and efficiency [58]. For the 2D PnCs considered here, FEM simulation with 2D plane strain approximation is performed for the PnC unit cell of Figure 2-5.a, using COMSOL Multiphysics solid mechanics module [76]. Floquet periodic boundary conditions (PBC) are set at the opposite edges of the 2D unit cell in both x and y directions, as shown in Figure 2-5.a. The wave vector $\vec{k}$ is used as a parameter for the PBC spatial periodicity. Eigenfrequency analysis is carried out sweeping $\vec{k}$ over the edges of the IBZ as in (2.42). The resulting dispersion relation for a PnC with the abovementioned dimensions in IBM 32 nm SOI is shown in Figure 2-6. The dispersion relation clearly

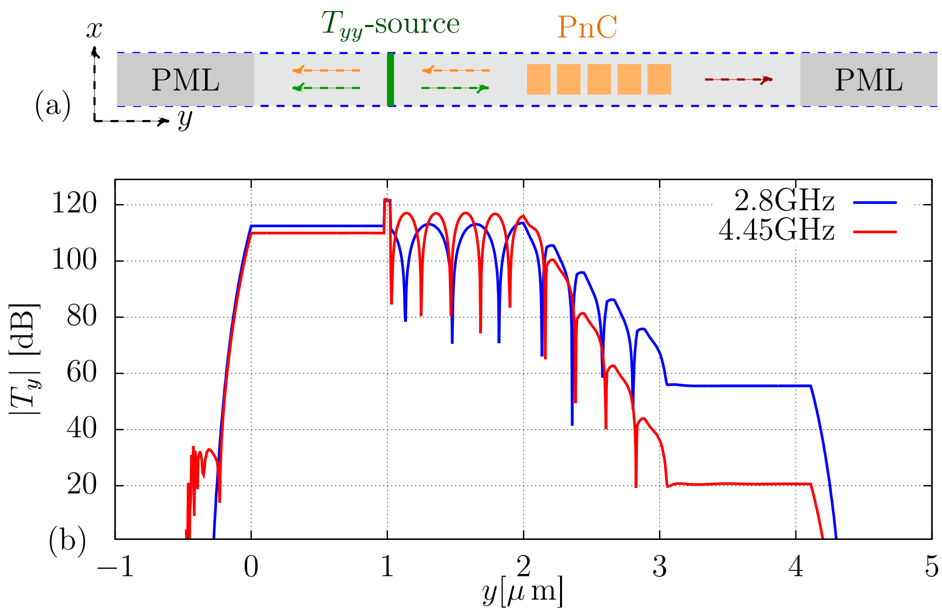


Figure 2-7: (a) FEM model for 5 layers of the PnC of Figure 2-5 (plotted horizontally), (b) Magnitude of the stress field $|T_{yy}|$ along the structure, showing: standing waves between the source and PnC, exponentially decaying waves in the PnC and transmitted waves after the PnC.

shows a complete, 85% fractional bandgap spanning 3.8GHz, centered at 4.45GHz (2.54GHz–6.35GHz). No waves are allowed to propagate in x or y directions in this bandgap.

2.2.2 Transmission Through BEOL PnC

Although a PnC implementation in CMOS is limited to the first few metal layers available in the technology, it will still achieve large reflectivity due to the high impedance contrast between the constituting materials. Consider the PnC of §2.2.1 in IBM 32 nm SOI where the impedance contrast is about 19× between SiCOH and Cu. This PnC is expected to achieve high reflectivity, even when using only the first 5 metal layers (subsequent metal layers have substantially different thicknesses, breaking the 2D periodicity assumption of this PnC). This is verified through FEM simulations of one PnC period in the x-direction as shown in Figure 2-7.a.

The vertical section of Figure 2-7.a is rotated horizontally for plotting convenience, such that the silicon substrate is on the left while the top of the CMOS stack is on the right. Periodic boundary conditions are enforced in the x-direction. A 1 MPa y-stress (T_{yy}) is applied from the wafer side, launching waves towards the PnC and the wafer bulk, where the latter is modeled as a perfectly matched layer (PML) [77]. A PML is also included on top (to the right in the rotated graph) of the PnC, to avoid any reflections back to the PnC.

Magnitude of the y-stress along the structure is shown in Figure 2-7.b. Standing waves are formed between the source and the PnC due to the reflections from the PnC. The magnitude of T_{yy} stress transmitted through the PnC is lower than the standing wave amplitude before the PnC by 57 dB and 89 dB at 2.8 GHz and 4.45 GHz, respectively. The strong evanescent decay of the stress wave inside the CMOS BEOL metal layers constitutes solid evidence for the efficiency of this PnC implementation, despite the limited number of metal layers available in the technology.

2.2.3 Z-Shape PnC Unit Cell

The unit cell of a different PnC design for IBM 32 nm SOI technology is shown in Figure 2-8.a. Copper vias are used together with the metalization to create a Z-like shape. The copper stripes have width of 125 nm and a spacing of 67 nm, both in x-direction. The only symmetry operation for this structure is inversion symmetry, and hence the IBZ consists of two coinciding rectangles in the first Brillouin Zone as shown in Figure 2-8.b. This is the case since time reversal symmetry sets $\omega(\vec{k}) = \omega(-\vec{k})$, and thus two coinciding rectangles are enough to determine the IBZ in the most general case, even with the absence of any structural symmetry for the unit cell.

The dispersion relation of this PnC is shown in Figure 2-9, which shows two major bandgaps of 1.13 GHz (2.23 GHz–3.36 GHz) and 1.07 GHz (7.13 GHz–8.2 GHz). This PnC can be advantageous for targeting higher frequencies, while still maintaining large enough metal lines width and separation for reliable manufacturing.

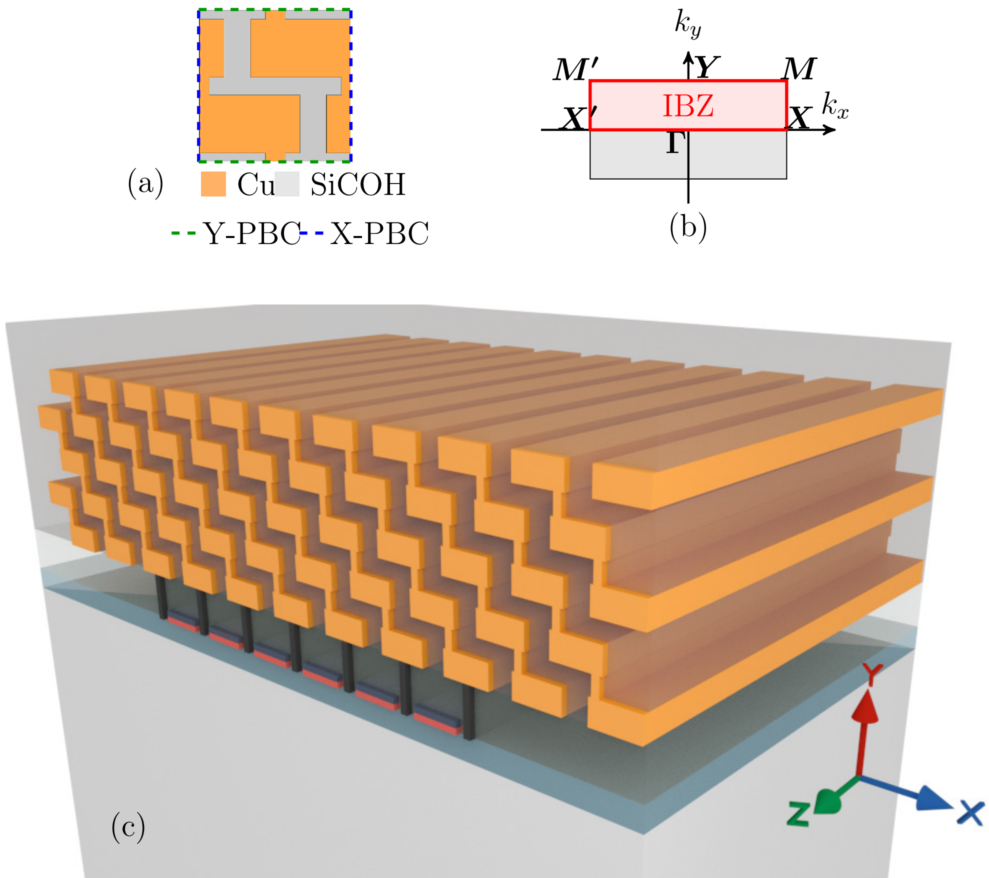


Figure 2-8: (a) A Z-shaped Unit cell for a PnC implemented in IBM32SOI technology BEOL, with only inversion symmetry. Blue and green dashed lines indicate the periodic boundary conditions applied in FEM simulation. (b) First Brillouin Zone and IBZ of the reciprocal lattice for the unit cell of (a). (c) 3D view of the full PnC on top of an RBT.

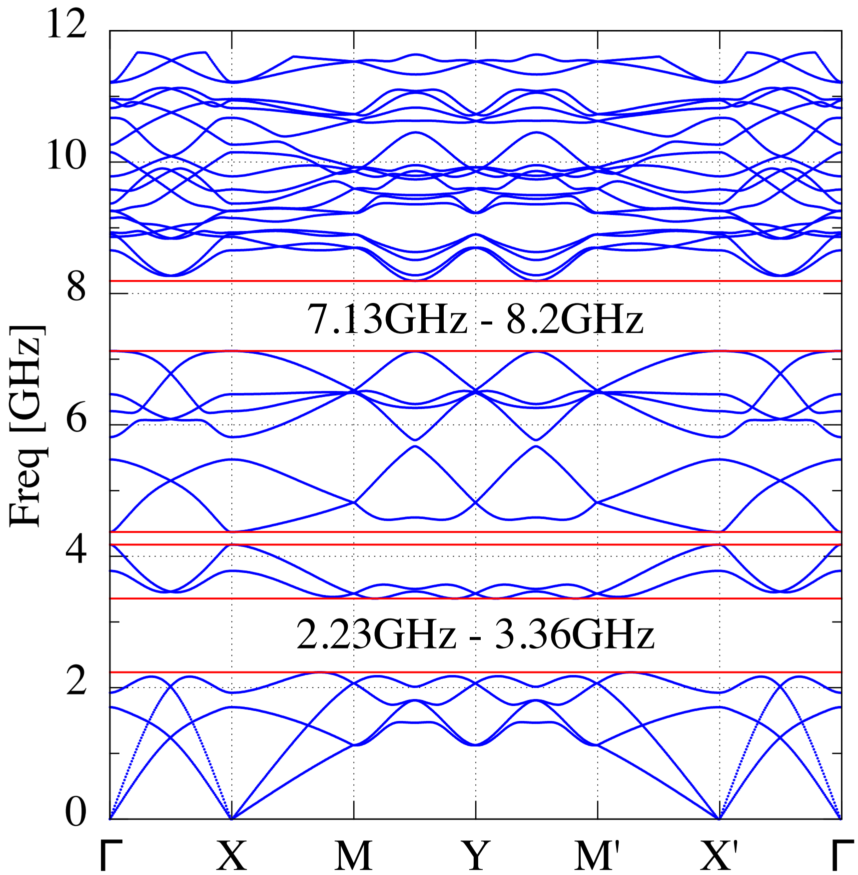


Figure 2-9: Dispersion relation for the PnC of Figure 2-8 implemented in IBM 32 nm SOI, showing bandgaps of 1.13 GHz (2.23 GHz–3.36 GHz) and 1.07 GHz (7.13 GHz–8.2 GHz).

2.2.4 PnCs in XFab $0.18\,\mu$m and IBM 130 nm

Another CMOS PnC implementation in XFab $0.18\,\mu$m BEOL layers is considered. The unit cell is shown in Figure 2-10 on the next page, while its dispersion relation is shown in Figure 2-11. This PnC is based on tungsten vias surrounded by SiO_2 inter-metal dielectric. As can be seen from Table 2.1 on page 60, these materials have large acoustic impedance contrast, making them ideal for constructing PnCs with large bandgaps as traditionally used in multiple phononic crystals implementations [55–57, 64]. With 850 nm tungsten vias separation, a 13% fractional bandgap (250 MHz from 1.80 GHz to 2.05 GHz) is obtained.

Another interesting feature of the PnC in XFab $0.18\,\mu m$ technology is the relatively small impedance mismatch between aluminum and SiO_2. This gives an extra degree of freedom for the aluminum routing traces without compromising the mechanical performance of the PnC. Furthermore, the lower wave velocity in SiO_2 compared to aluminum, creates the possibility for index guiding between the aluminum layers, which further improves the performance of such PnCs.

PnC design in IBM $0.13\,\mu$m technology with a unit cell similar to that of Figure 2-5.a is also considered. In this particular technology, the routing layers are formed from copper metallization surrounded by SiO_2 isolation. As can be inferred from Table 2.1 on page 60, the contrast in acoustic impedance between these materials is not as strong as the SiCOH/Cu case; hence a smaller bandgap is to be expected for this design.

The dispersion relation for a typical PnC in this technology is shown in Figure 2-12. This PnC shows a bandgap of 310MHz (7% fractional bandgap) for 300 nm-wide metal lines with 200 nm separation. Such a bandgap was found to be among the highest attainable in this technology, which is expected due to the smaller contrast in acoustic impedance. Moreover, the DRC restrictions force the metal segments to be far apart, reducing the resulting bandgap when compared to other technologies.

This technology has layer thickness variations that significantly exceed its fractional bandgap. The performance of this PnC will be sensitive to process variations,

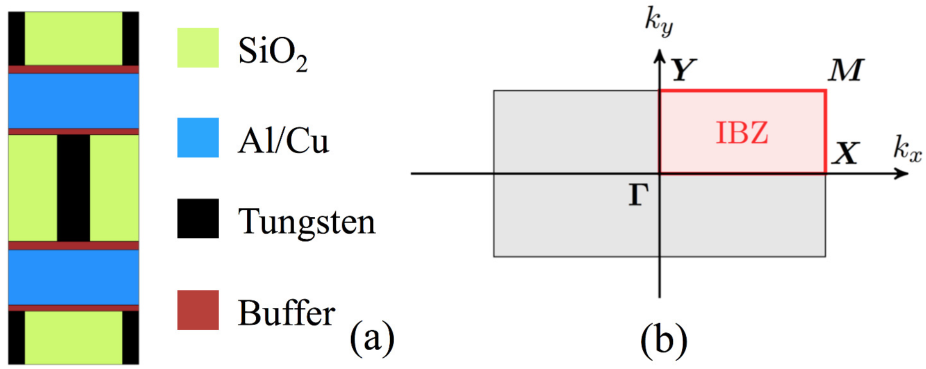


Figure 2-10: (a) Unit cell for a PnC implemented in XFab 0.18 μm technology BEOL. (b) First Brillouin Zone and IBZ of the reciprocal lattice for the unit cell of (a).

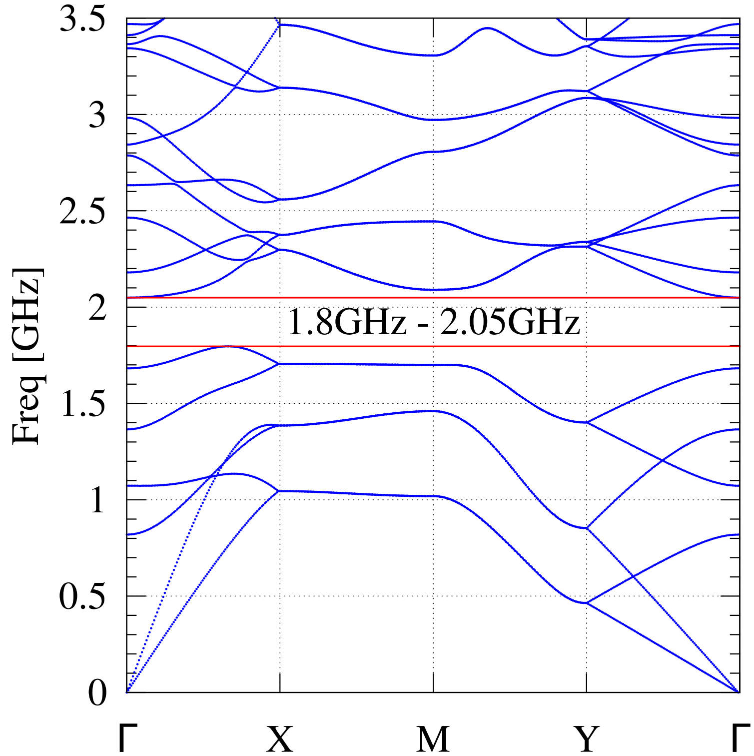


Figure 2-11: Dispersion relation for the PnC of Figure 2-10 implemented in XFab 0.18 μm bulk CMOS technology, showing a bandgap of 250 MHz (1.80 GHz–2.05 GHz).

significantly limiting its usability for practical implementations.

Table 2.2 compares the performance of all the PnCs considered in this chapter. It is clear that the design of Figure 2-5 on page 61 shows the largest bandgap and will be the focus of the further studies included here.

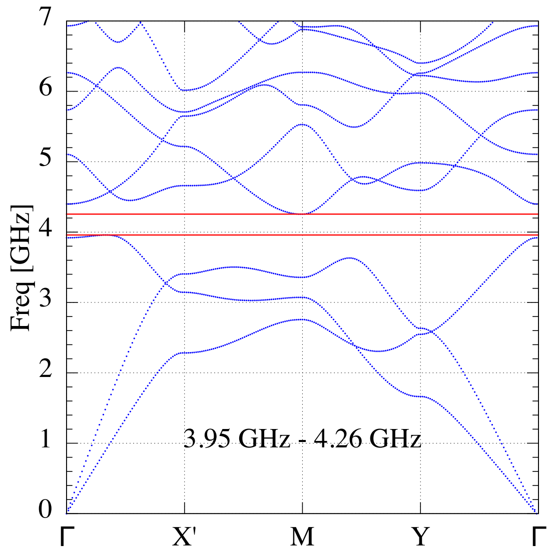


Figure 2-12: Dispersion relation for a typical PnC in IBM 0.13 μm technology, showing a bandgap of 310 MHz (3.95 GHz–4.26 GHz).

Table 2.2: Performance summary of different CMOS PnC considered in this study.

Technology	IBM			XFab
	32 nm SOI	32 nm SOI	0.13 μm	0.18 μm
Unit Cell	Fig. 2-5(a)	Fig. 2-8(a)	Fig. 2-5(a)	Fig. 2-10(a)
$\omega - \vec{k}$	Fig. 2-6	Fig. 2-9	Fig. 2-12	Fig. 2-12
Dielectric	low-κ SiCOH		SiO_2	SiO_2
Scatterers	Cu	Cu	Cu	W
Z_{11} ratio	19	19	3	7.5
Z_{44} ratio	15.9	15.9	2.3	6.6
Bandgap Width [GHz]	3.8	1.13 1.07	0.3	0.72 0.46
Bandgap Center [GHz]	4.45	2.8 7.67	4.1	1.79 2.92

2.3 CMOS Process Variations and PnCs

Process variations inherent to commercial CMOS processes, with their random nature, represent a non-trivial challenge for PnC implementation therein. Variations in BEOL layer thicknesses, metal widths and lithographic misalignment are notable examples. Furthermore, material properties may change depending on conditions during deposition or growth as well as the residual stress. Equation (2.33) suggests that uniform isotropic scaling of the PnC structure will result in linear scaling of the PnC dispersion relation. In case of uniform anisotropic scaling, dispersion relation variations are expected to be sublinear. However, the aforementioned CMOS process variations are far from uniform scaling. Owing to their stochastic nature, these variations induce a *mismatch* between the different PnC layers.

2.3.1 Perturbation Theory and Stored Energy

Process variations in general can be considered as perturbations of an ideal PnC structure. Perturbation theory is a useful framework to analyze the effects of such process variations. Small perturbations of the structure geometry induce a frequency shift $\Delta\omega$ that is given by [70]

$$\frac{\Delta\omega}{\omega^0} = \frac{1}{2}\left(\frac{\Delta U}{U^0} - \frac{\Delta K}{K^0}\right), \tag{2.43}$$

where U and K are the stored potential and kinetic energy, respectively. The superscript 0 denotes the unperturbed modes. $\Delta U = U - U^0$ and $\Delta K = K - U^0$ represent the change in the stored energies due to perturbation.

Thickness variations and dimensions mismatch actually result in the *shifting* of material boundaries. Based on (2.43), it is expected that modes with significant energy storage near the perturbation location (in this case the shifted boundary), will be more likely to experience larger frequency shifts. Although such variations with shifting boundaries can *not* be considered *small* perturbations [78], the former reasoning concerning stored energy remains valid. FEM simulation will always be

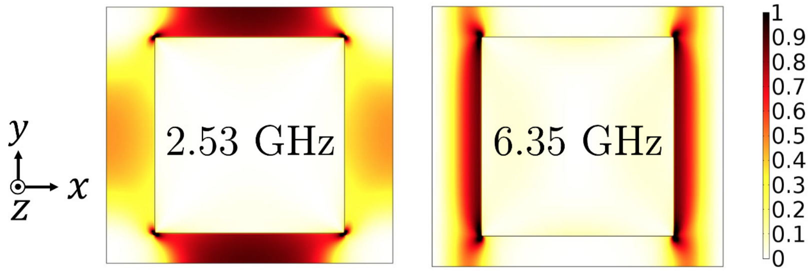


Figure 2-13: Normalized stored strain energy distribution for the PnC unit cell of Figure 2-5.a, plotted for modes at the edges of the PnC bandgap with $k_x = 0$ and $k_y = 0$.

necessary to obtain accurate numerical results on the effects of each variation type.

The actual effect of process variations strongly depends on the actual PnC geometry and material composition. The IBM 32 nm SOI PnC of Figure 2-5.a will be used to study the effects of different CMOS process variations on PnC performance. Figure 2-13 shows the normalized strain energy distribution within the unit cell of the aforementioned PnC, for the modes just above and just below the bandgap, with $k_x = 0$ and $k_y = 0$. The effect of layer thickness and metal width variations on these modes strongly depend on the distribution of the strain energy.

2.3.2 Simulation Framework Based on PnC Transmission

When considering process variations and mismatch, the phononic crystal is no longer a perfectly periodic medium in all directions: such variations present themselves as defects in the crystal structure. The effect of individual metal and dielectric layers perturbations cannot be studied in a PnC unit cell with periodic boundary conditions. For this reason, the 2D FEM model of Figure 2-7.a is used in a frequency domain simulation to find the transmission coefficient $\boldsymbol{T}$ of the PnC, defined as

$$|\boldsymbol{T}| = \left|\frac{T_t}{T_i}\right| \tag{2.44}$$

where T_t and T_i are the transmitted and incident stresses respectively. This problem can be approached as a transmission line problem by exploiting the similarity between elastic and electro-magnetic waves [68]. The transmitted stress T_t is found directly from the amplitude of the travelling wave transmitted through the PnC. The incident stress T_i is found from the standing wave formed before the PnC, by first finding the standing wave ratio (SWR)

$$\mathrm{SWR} = \left|\frac{T_{\max}}{T_{\min}}\right| \tag{2.45}$$

where $T_{\max}$ and $T_{\min}$ are the maximum and minimum stresses of the standing wave, respectively. The reflection coefficient $\mathbf{\Gamma}$ and hence the incident stress are then found as

$$|\mathbf{\Gamma}| = \left|\frac{T_i}{T_r}\right| = \frac{\mathrm{SWR}-1}{\mathrm{SWR}+1} \tag{2.46}$$

$$T_i = \frac{T_{\max}}{1+|\mathbf{\Gamma}|} \tag{2.47}$$

where T_r is the reflected stress.

2.3.3 Mismatch in Horizontal Dimensions (Metal Width)

First, the width of individual metal layers is assumed to increase, emulating one of the possible mismatches in the PnC. From the energy distribution of Figure 2-13, large stored energy is localized around the metal line's side walls. It is thus expected that the mode at 6.35 GHz, and hence the upper edge of the bandgap, will be more susceptible to this perturbation.

Figure 2-14 compares the transmission coefficient of an ideal PnC to one that is subject to such mismatch. The upper bandgap edge is significantly affected by such variation in comparison to the lower edge as predicted. Aside from the bandgap edge shift, no spurious transmission is observed for such variation. With sufficiently large bandgap, this type of variation is not a major concern for resonator designs.

2.3.4 Mismatch in Vertical Dimensions

Next, the thickness of the metal and via layers are perturbed by 20% of their nominal thickness. Perturbations to both individual layers and multiple layers are tested. The energy distribution of Figure 2-13 on the preceding page shows large stored energy

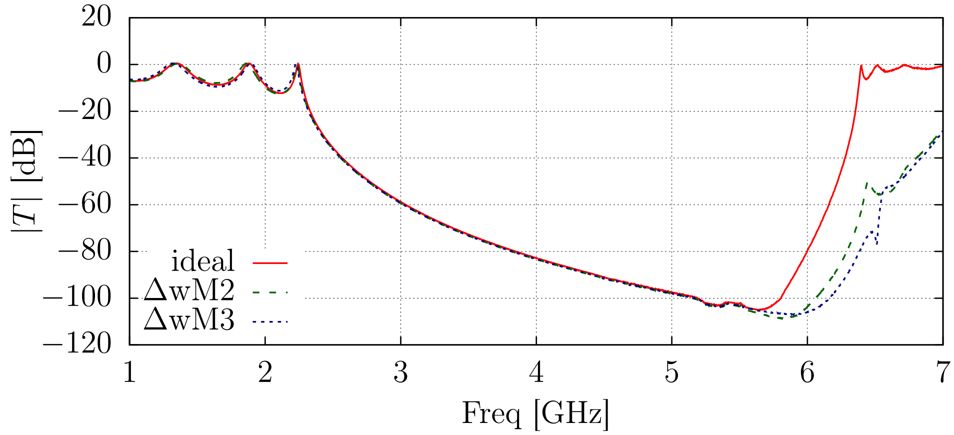


Figure 2-14: Transmission coefficient through PnC of Figure 2-5.a with variation in the metal lines' width. ΔwM2 and ΔwM3 indicate +20 nm individual changes in the width in x-direction of metal 2 and 3 respectively.

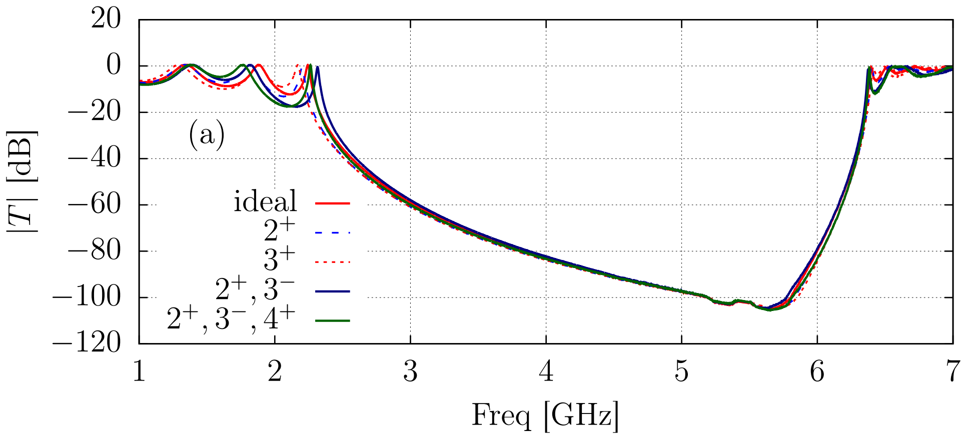
concentration around the metal lines' top and bottom surfaces. It is thus expected that the lower bandgap edge mode at 2.53 GHz will suffer larger frequency shifts than the upper bandgap edge.

Figure 2-15 and Figure 2-16 show the transmission coefficients for PnCs with such perturbation. Both figures clearly highlight the larger shift in the lower bandgap edges in agreement with the prediction. Maximum shifts of 74 MHz and 142 MHz are observed in the lower bandgap edges for the metal layer and via thickness perturbation, respectively. It is important to notice that for a PnC with a small number of layers (in this case 5 layers), the PnC performance is symmetric in terms of perturbation: perturbing the M4 layer is equivalent to perturbing the M2 layer. Again, no spurious transmission is observed for this type of variation, and designs with large bandgaps are less affected by this process variation.

2.3.5 Lithographic Misalignment

Lithographic misalignment between the different metal layers is also considered as a major process non-ideality. A 20 nm misalignment between the metal layers is assumed. Figure 2-17 shows the resulting transmission both for individual layer misalignment along with 2- and 3-layer simultaneous misalignment.

First, the upper edge of the bandgap is affected by such misalignment, as expected



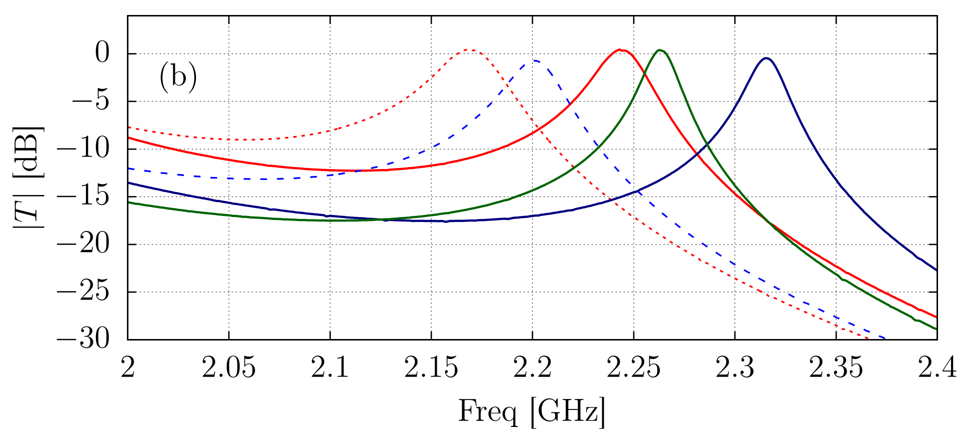
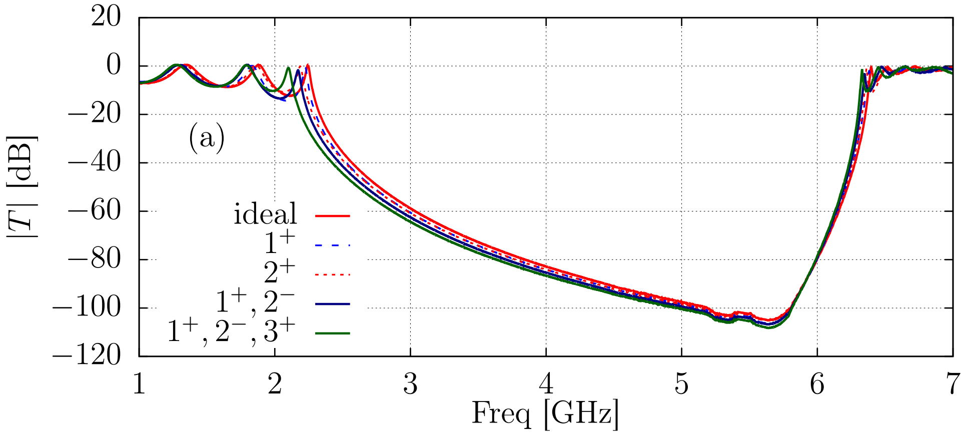


Figure 2-15: (a) Effect of different metal layer thickness variation on the transmission coefficient through the PnC of Figure 2-5. $n^{\pm}$ indicate a ±20% change in the thickness of the n^{th} layer; (b) Zoomed in version of (a) around the lower PnC band edge.



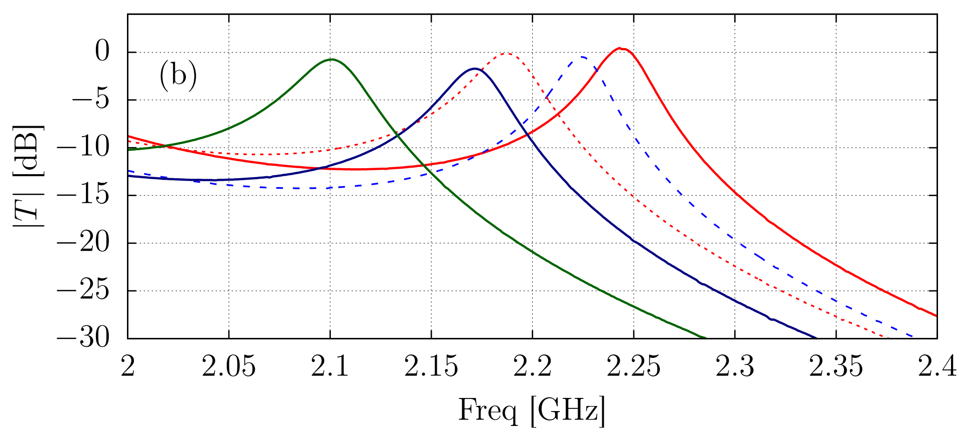


Figure 2-16: (a) Effect of different via layer thickness variation on the transmission coefficient through the PnC of Figure 2-5. $n^{\pm}$ indicate a $\pm 20\%$ change in the thickness of the n^{th} via layer; (b) Zoomed in version of (a) around the lower PnC band edge.

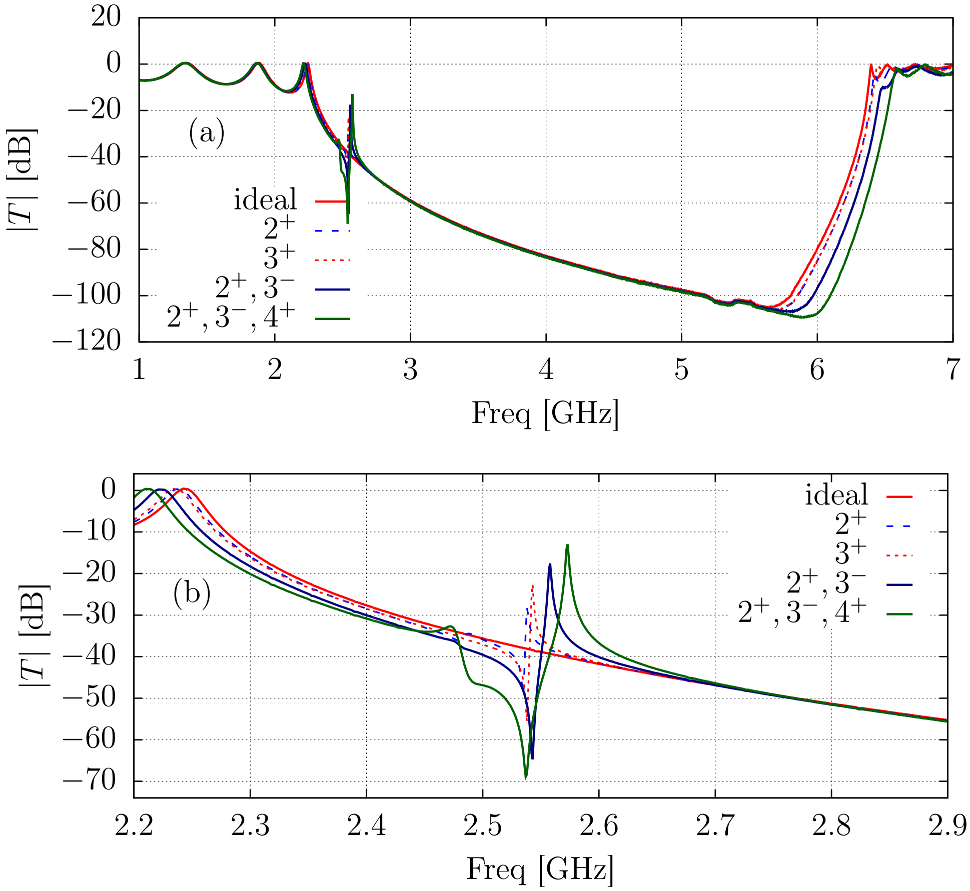


Figure 2-17: (a) Effect of different metal layers lithographic misalignment on the transmission coefficient through the PnC of Figure 2-5. $n^{\pm}$ indicate a $\pm 20\%$ misalignment in the x-direction for the n^{th} metal layer; (b) Zoomed in version of (a) around the spurious transmission.

from the stored energy argument. More importantly, a spurious transmission mode is created around 2.55 GHz. This can be explained as the misalignment creating a line defect resonance cavity in the PnC, which introduces a defect mode in the PnC bandgap. Evanescent modes in the PnC couple energy to this cavity, creating the transmission spur [66]. The sharp dip in the transmission indicates a spurious change in the reflectivity and hence the acoustical impedance of the PnC changes. Such spurious reflections have the potential to create spurious modes in the RBT frequency response if the corresponding stress-strain field distribution can be driven and sensed by the RBT transducers.

2.4 Conclusion

BEOL materials in CMOS can be used to create PnCs with large bandgaps, useful for energy confinement in high-Q unreleased CMOS-MEMS resonators. Simple periodic metal lines implemented in the BEOL of IBM 32 nm SOI technology, form a PnC with a large 85% bandgap as shown in Figure 2-5 on page 61. Different PnC designs have also been considered, including tungsten-SiO_2 PnC in XFab 0.18 μm. Table 2.2 on page 69 compares the performance of the different PnCs considered in this chapter.

The effect of CMOS process variations on the IBM 32 nm SOI PnC of Figure 2-5 has been studied. Typical process variation and mismatch mostly affect the PnC characteristics near its bandgap edges. In case of PnCs with wide bandgaps, the PnC characteristics are negligibly affected near the center of the bandgap. Achieving wide bandgaps through the selection of materials and dimensions is thus necessary for PnC implementation in CMOS for several reasons:

- The decay rate of evanescent waves in the PnC depends on the frequency separation from bandgap edges; hence wider bandgap ensures faster decay;
- Faster wave decay enables high reflectivity with a small number of layers (a common limitation of CMOS processes); and
- The wide bandgap provides larger tolerance for random process variations and different mismatch.

This explains why the PnC implemented in IBM 0.13 μm (Figure 2-12), with a fractional bandgap of 7%, is highly unreliable for practical implementations, as discussed in section 2.2.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 3

PnC Waveguides, Perturbations and Adiabatic Transitions

The CMOS BEOL PnCs presented in the previous chapter can be used for a myriad of applications. The focus in this thesis is aimed at acoustic wave confinement to form unreleased CMOS-MEMS resonant body transistors (RBTs). The RBT cavity will naturally be located around the CMOS FEOL layers to take advantage of MOSCAP electrostatic driving and active piezoresistive FET sensing. With the PnC existing only in the BEOL layers, it is tempting to think that it is only useful for providing reflections from the top surface of the RBT cavity. However, by virtue of the PnC periodicity and the higher sound velocity in the bulk silicon, an acoustic waveguide is created between the BEOL PnC and the bulk wafer. For a specific range of $\vec{k}$, waves are reflected from the bulk silicon wafer, a phenomenon similar to index guiding in photonics. Waveguides are key components in a resonance cavity.

This chapter focuses on the study of periodic acoustic waveguides based on BEOL PnCs [79, 80]. The effects of perturbations on the waveguide characteristics are also considered.

3.1 PnC-Based Phononic Waveguides

Consider a simple structure in IBM 32 nm SOI technology formed from the BEOL PnC (Figure 2-5 on page 61) on top of unpatterned FEOL layers (excluding contacts) and the bulk wafer (this is similar to Figure 2-5.c without the FEOL MOSFETs). Such structure is periodic in the x-direction with the same period as the PnC. A vertical section corresponding to a single horizontal period in the x-direction with x-periodic boundary conditions (x-PBC) is considered as shown in Figure 3-1.a.

3.1.1 Waveguide Dispersion Relation

2D FEM simulations are used to find the dispersion relation of this periodic structure. Owing to its periodicity in the x-direction only (a discrete translational symmetry), the k_x component of the $\vec{k}$-vector is conserved (in space and time) throughout the entire structure as discussed in §2.1.3. For this reason, only k_x is used as a parameter for the dispersion relation calculation as shown in Figure 3-1.b. Different types of modes can be observed in this dispersion relation [71]:

- *PnC propagating modes*: Modes that can propagate into the PnC and form continuous bands corresponding the PnC bands (bands 1 and 2).

- *Bulk propagating modes*: Plane waves propagating in the bulk wafer. Their dispersion relation is simply given by

$$\omega = c|\vec{k}| = c\sqrt{k_x^2 + k_y^2}, \tag{3.1}$$

  where c is the wave velocity (being longitudinal or shear waves). Thus, for a given k_x, the allowed modes in the bulk form a continuum $\omega > c\,k_x$, referred to as the *sound-cone* and bounded by the *sound-line* $\omega = c\,k_x$.

- *Bulk evanescent modes*: Modes below the sound-line, with $\omega < c\,k_x$ and imaginary k_y. They decay exponentially in the bulk in the y-direction. This phenomenon corresponds to index guiding in photonic waveguides: a generalization of the total internal reflection obtained from Snell's law with incidence angle bigger than the critical angle.

- *PnC evanescent modes*: Located in the PnC bandgap, they decay exponentially in the PnC. These modes are not shown in PnCs' dispersion relations since they do not exist in infinite, perfectly-periodic structures.

- *Localized modes*: Modes both below the sound-cone and in the PnC bandgap. Such modes show evanescent exponential decay in both the PnC and bulk.

The localized modes are spatially confined in the FEOL layers between the PnC and the bulk silicon wafer. This confinement results in a discrete spectrum of modes as in the case of photonic waveguides and quantum wells [69, 71]. Without further constraints in the lateral dimensions, wave propagation is allowed between the PnC and bulk wafer in the $x-z$ plane, forming a horizontal waveguide. Such waveguiding provides perfect acoustic confinement for the RBT. The full resonator structure imposes additional constraints to confine the waves in the horizontal direction, creating a resonance cavity.

It is important to notice that the buried oxide (BOX) layer of the SOI wafer plays no role in the formation of this horizontal acoustic waveguide. The waveguide is formed mainly because the sound velocity in the bulk is larger than that in the surface and the fact that these modes are inside the PnC bandgap and cannot propagate upwards. Thus PnC-based horizontal acoustic waveguides are not exclusively limited in any way to SOI technologies: the concept is readily applicable in CMOS bulk technologies as well.

3.1.2 Engineering the Dispersion Relation

Based on (2.33), isotropic scaling of the entire structure will result in a proportional scaling of the eigenfrequencies thereof. This scaling can be readily used to engineer the dispersion relation of the phononic waveguides. However, since in CMOS technologies the designer can only control the horizontal dimensions, the resulting scaling is not isotropic. The corresponding scaling of eigenfrequencies becomes sublinear.

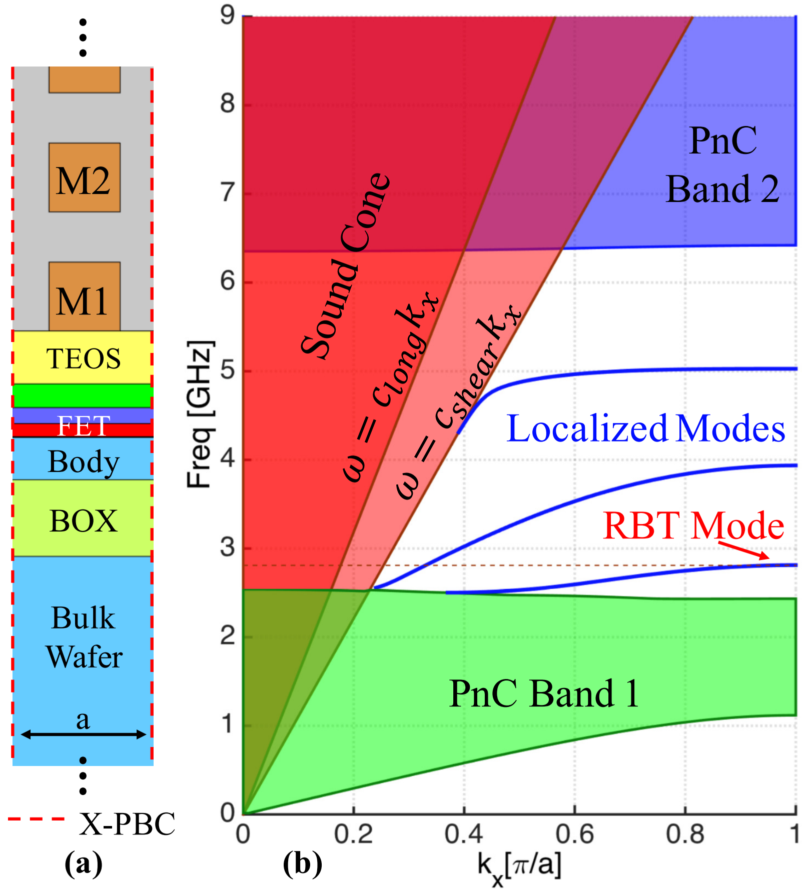


Figure 3-1: (a) Vertical section of the PnC of Figure 2-5.a showing FEOL layers and the wafer bulk with x-periodic boundary conditions (X-PBC). (b) Dispersion relation from FEM simulation showing PnC bands, sound cone and localized modes.

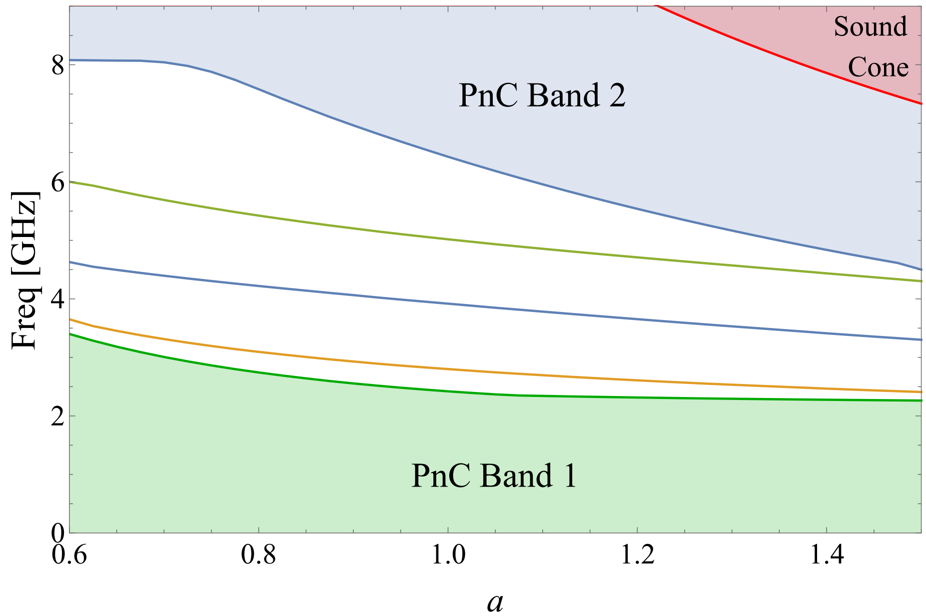


Figure 3-2: Evolution of the eigenfrequencies at $k_x = \pi/a$ for the waveguide of Figure 3-1 as function of the period a. The period is normalized to that of Figure 3-1 with uniform scaling in x-direction.

Figure 3-2 shows the evolution of the waveguide eigenmodes as a function of the normalized waveguide period, resulting from geometry scaling in the x-direction. Although the eigenfrequency scaling is sublinear with dimensions, it is still sufficient to provide a degree of freedom in the phononic waveguides design. Actual waveguide design may require separate tuning for the metal width and separation of the phononic crystal to further optimize the desired mode shape (stress distribution) in the waveguide as well as its dispersion characteristics.

Being able to engineer the dispersion relation of the waveguide is a useful capability. The foremost implication is that it provides a control over the resonance frequency of the resulting RBT cavity. Moreover, in the same resonator, waveguides with different dispersion relation can be used to create a distinct separation between a main cavity and a termination section. This design strategy will be further explored in chapter 4.

3.2 The First PnC-Based RBT

The primary application for phononic waveguides considered in this work is the realization of unreleased RBTs in CMOS. A CMOS-RBT has been previously demonstrated in [36] that relied solely on acoustic Bragg reflectors (ABRs) for acoustic confinement [34]. While this was the first demonstration of unreleased CMOS-MEMS resonators, this resulted in a low quality factor ($Q < 30$) due to the small solid angle subtended by the ABRs and hence increased radiation losses through both the substrate and the BEOL layers. The phononic waveguides presented in the previous section are ideal candidates to achieve almost perfect vertical energy confinement. The first PnC-based RBT was demonstrated in IBM 32 nm SOI technology* [79]. The proposed phononic waveguides help explain the major performance enhancement obtained for this RBT generation.

3.2.1 First PnC-Based RBT Structure

Figure 3-3 shows a cross-section SEM micrograph of the first PnC RBT where the phononic crystal on top of the bulk wafer is clearly highlighted. The RBT uses the foundry-provided analog nFETs both for sensing and electrostatic driving. This allows the RBT to benefit from the ultra-thin, high quality gate dielectric for efficient driving by internal dielectric transduction, as well as the highly reliable, high-f_T nFETs for active transistor sensing [32, 35, 36].

The main RBT mode is a localized waveguide mode with evanescent decay in both the PnC and the bulk wafer as discussed in §3.1. Large tungsten bulk-ties have been used to terminate the RBT waveguide cavity horizontally. Moreover, the drain-side contacts from the MOSCAP were removed to reduce internal scattering.

The RBT is driven by two nFETs used as MOS capacitors (MOSCAPs), while a single sensing nFET was included at the center of the cavity between the two driving MOSCAPs. The device layout is shown in Figure 3-4. This implementation uses

*This project began as a collaboration with R. Marathe, who helped with FEM simulations of the RBT structure as well as early layout efforts.

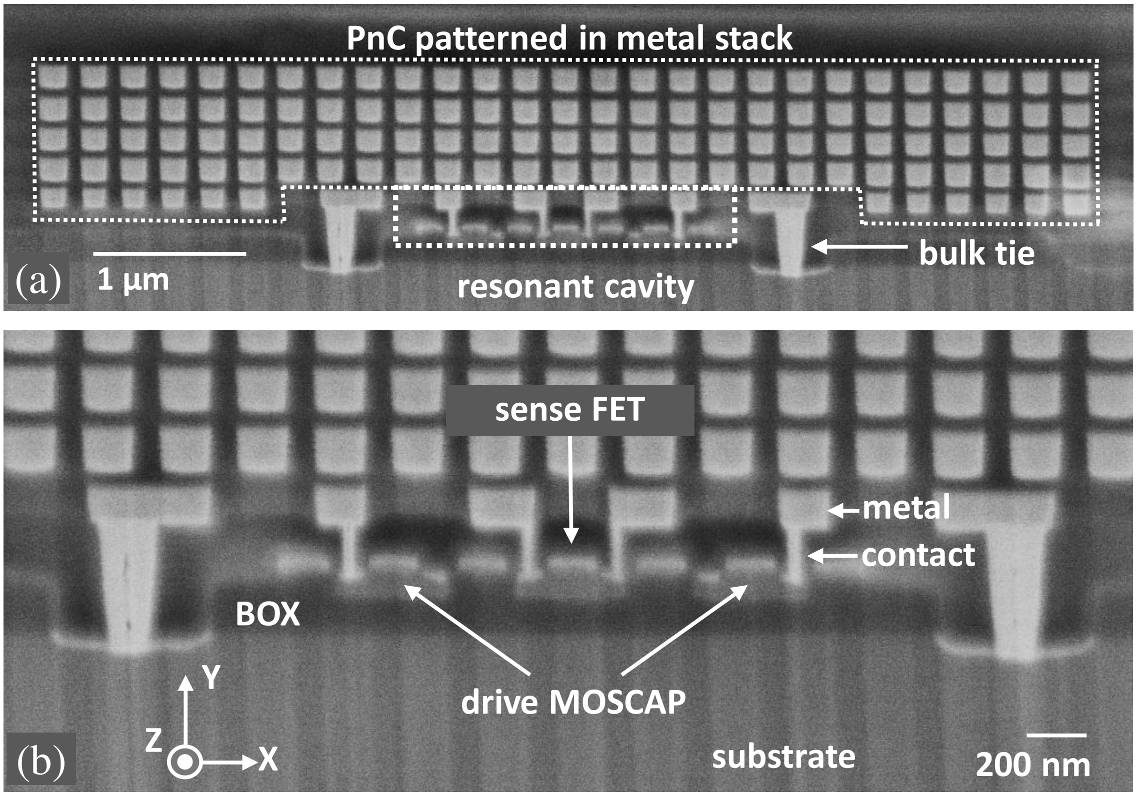


Figure 3-3: (a) SEM micrograph of an unreleased RBT implemented in IBM 32 nm SOI technology based on a PnC waveguide. (b) A detailed micrograph of the resonant cavity showing the driving MOS-Caps along with the sensing FET.

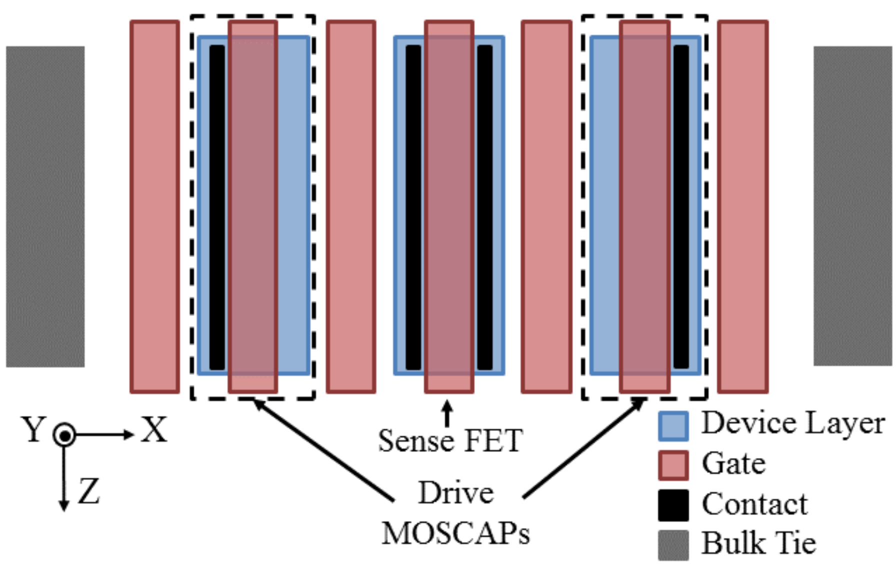


Figure 3-4: Top view of PnC waveguide resonator of Figure 3-3 layout showing drive and sense transistors, gates, modified contacts, first metal level and bulk ties. Metal layers are excluded for clarity of underlying structure.

nFETs with 160 nm gate length, 2 μm width and gate-to-gate pitch of 290 nm.

The RBT structure is uniform in the out-of-plane z-direction. Long, rectangular, wall-like vias spanning the entire width of the FET were used as shown in Figure 3-4. Maintaining uniformity along z-direction (with sufficient structure width) is important to reduce scattering, as will be discussed in detail in §3.4. A floating-body nFET was chosen for sensing over a body-contacted one, in order to reduce cavity perturbations in the z-direction. Moreover, the drive and sense gate routing was limited to the first metal layer to avoid perturbing the PnC in the higher metal layers.

Due to the strong mismatch between the acoustic impedance of the routing metals and dielectrics, the electrical routing and connections to the device can potentially interfere with its mechanical operation. This represents one of the challenges of the structural design of such unreleased resonators.

3.2.2 RBT Driving and Sensing FETs

The spatial configuration of the driving FETs favors a distribution of k_x values near $k_x = \pi/a$. Specific waveguide modes from the dispersion relation of Figure 3-1.b are excited. These modes become the resonator RBT mode. As will be demonstrated in §3.4, driving near $k_x = \pi/a$ is beneficial for reducing scattering to the bulk wafer sound cone, enabling larger quality factors. The driving nFETs are operated in strong inversion, where their behavior closely matches that of a parallel plate capacitor in terms of the generated stress.

The advantage of FET sensing in CMOS RBTs over other transduction mechanisms is discussed in §1. The structural integrity of the FETs must be maintained to avoid compromising their performance. For this reason, the foundry-included dummy gates for the sensing FET were left unmodified and double as dummy gates for the driving MOSCAPs as well. This guarantees transistor channel length uniformity and preserves the structure of the stress liners, which is crucial for setting the FET channel mobility and threshold voltage.

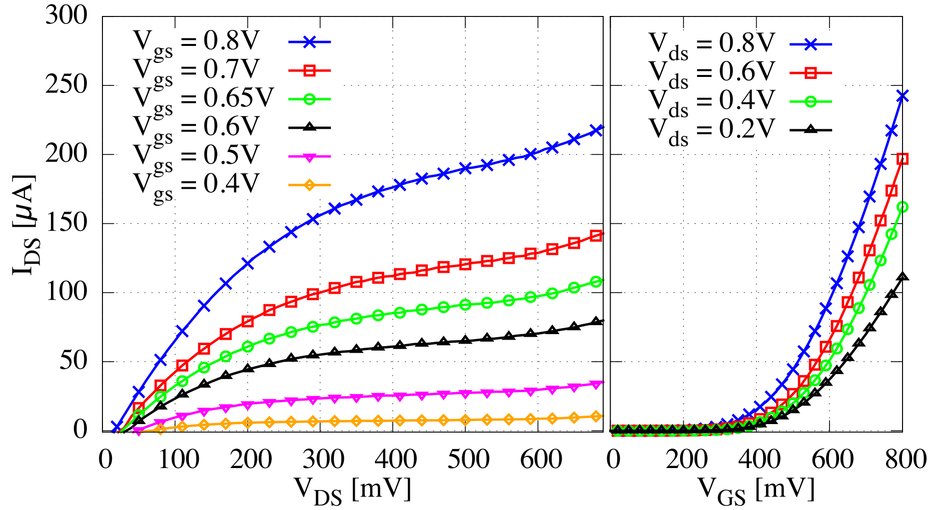


Figure 3-5: (left) $I_{DS} - V_{DS}$ characteristics and (right) $I_{DS} - V_{GS}$ of the resonator's sensing FET showing expected transistor response. In operation, the sensing FET was biased at $V_{DS} = 0.6\,\text{V}$ and $V_{GS} = 0.65\,\text{V}$.

3.2.3 Characterization of the PnC RBTs

The first PnC RBTs were fabricated in the standard IBM 32 nm SOI process without any additional post-processing of any kind. The RBT of Figure 3-3 occupies an area of $5\,\mu m \times 7\,\mu m$. The measured DC characteristics of the sensing FET are shown in Figure 3-5, which closely match the projected performance based on the provided IBM design kit. This verifies the structural integrity of the fabricated MOSFETs. For normal RF operation, the driving MOSCAP gates are biased in inversion with $V_A = 1\,V$ while the sensing transistor is biased in saturation* with a grounded source, while the drain and gate voltages are set for $V_{DS} = 0.6\,V$ and $V_{GS} = 0.65\,V$, respectively. The drain current in the sensing FET in this case is $I_{DS} = 95\,\mu A$, well within the electromigration limits of the RBT routing metals. The saturation regime has the benefit of maximizing the output resistance of the FET, which greatly simplifies the subsequent design of current readout circuits.

A Cascade PMC200 RF probe system was used to test the presented RBTs. TRL calibration up to the GSG probe tips was carried out at room temperature with -10 dBm input power and 2 kHz IF bandwidth with 50 averaging traces using an Ag-

* Subsequent analysis shows that linear operation is much better in terms of maximizing the channel mobility sensitivity to the strain.

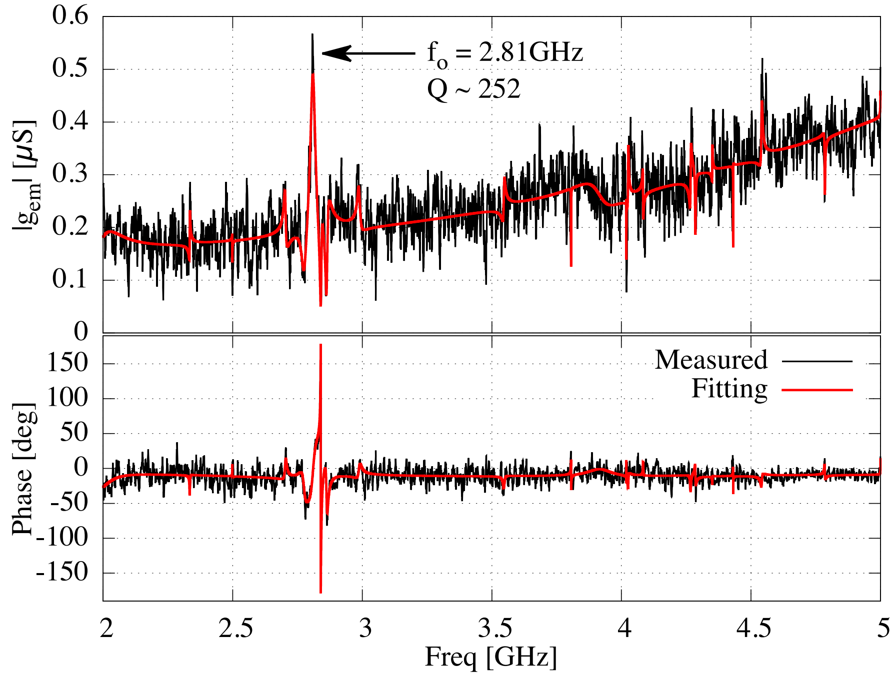


Figure 3-6: Measured electromechanical transconductance g_{em} of the IBM 32nm SOI RBT of Figure 3-3.

ilent PNA N5225A. The overall electromechanical transconductance g_{em} is calculated from the de-embedded Y-parameters, per standard π-model for MOSFETs, as

$$g_{em} = i_{out}/v_{in} = Y_{21} - Y_{12}. \tag{3.2}$$

Electrical parasitics of the RBT were de-embedded using its own response at $V_A = 0\,V$ on the driving MOSCAP, corresponding to the resonator "Off" state to suppress the mechanical mode. The measured g_{em} frequency response of a 2.8 GHz resonator is presented in Figure 3-6. An 11-point running average smoothing filter is applied to smooth the measured data. This smoothing is a highly conservative operation that may reduce the quality factor of the resonator peak. After smoothing, a rational transfer function fitting [81] with 24 poles was used to extract a Q of 252 at the 2.8 GHz resonance peak, leading to an $f \times Q = 7 \times 10^{11}$. By virtue of the superior confinement achieved by the PnC, the presented RBT design shows an 8× improvement in Q (2.5× in $f \times Q$) over the previous-generation CMOS-integrated RBTs in [36] with only 2× increase in the overall footprint. At such frequencies, the dominant phonon relaxation

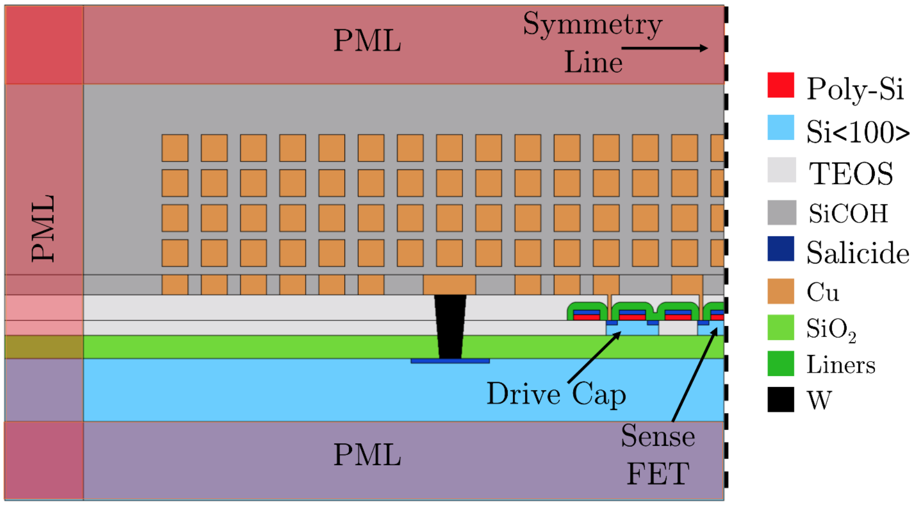


Figure 3-7: FEM model for the RBT structure, showing PMLs (scaled for better presentation) and symmetry boundary conditions.

mechanism is the Landau-Rumer regime, where the $f \times Q$ product increases linearly with frequency [82]. Since both resonators' $f \times Q$ product is significantly lower than the expected theoretical limits of most materials, the quality factor is limited by radiation loss instead of phonon relaxation. For this reason, the PnC RBT shows a major improvement in the quality factor Q. The PnC RBT also shows a wider spurious-free spectral range extending up to 4.5 GHz.

3.2.4 RBT FEM Simulation

2D plane-strain-approximation frequency domain FEM simulation is carried-out for a cross section of the RBT structure in Figure 3-3. The FEM model is shown in Figure 3-7, where only half of the RBT structure is considered with symmetry boundary conditions. The COMSOL Multiphysics solid-mechanics module is used [76]. The FEM model was calibrated with the exact fabricated dimensions as found from the SEM micrograph (including individual PnC layer thicknesses). Perfectly matched layers (PML) surround the RBT to prevent reflections to the resonance cavity, hence approximating the radiation losses [77, 83]. Stress load is applied to the driving MOSCAPs, corresponding to the actual electrostatic stress generated by the transducer.

The resulting average stresses in the sensing transistor channel area are shown in Figure 3-8. The measurement results are found to be in good agreement with the simulation predictions for the resonance frequency. The spurious mode at 4.6 GHz corresponds to the high-frequency mode of the simulation but is not distinguishable above the feedthrough floor and has intrinsically lower Q. The measured Q was expected to be lower than the simulated Q value of 962 for two reasons:

(a) The inter-metal dielectric is a porous SiCOH (pSiCOH) ultralow-κ dielectric [84] and is expected to be a significant source of viscoelastic damping [85]. However, relevant material parameters could not be disclosed by the foundry or were unknown, and hence viscoelastic damping was not included in the simulation, which only considers lateral radiation losses from the cavity (in the $x-y$ plane).

(b) The simulation is a 2D plane strain approximation and does not account for scattering due to the finite cavity depth as well as its abrupt termination as discussed in section 3.4.

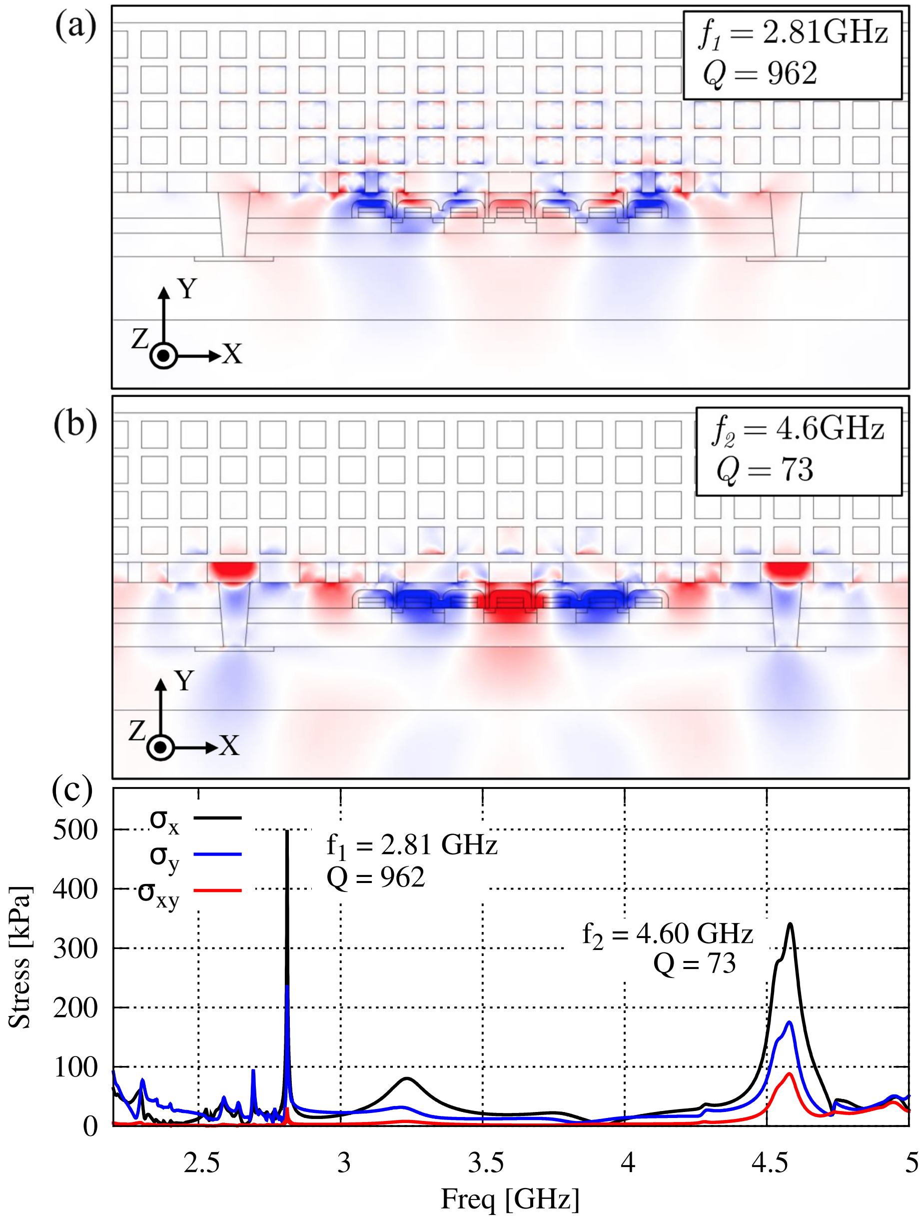


Figure 3-8: (a) X-stress at 2.81 GHz in 2D PnC-RBT structure showing Q of 962; (b) X-stress in resonant structure at 4.6 GHz with Q of 73; (c) Frequency response from 2D FEM COMSOL simulation of the RBT, showing the average stresses in the channel area with resonances at 2.81 GHz (Q of 903) and 4.5 GHz.

3.3 Perturbations in Phononic Waveguides

We now study the effect of different imperfections and perturbations on the waveguide structure. While it is always possible to estimate the performance of the perturbed waveguide using ab-initio techniques, perturbation theory is an efficient tool to assess waveguide performance changes for small perturbations. In this section, an operator-theoretic approach for the perturbation theory in phononic waveguides is presented. The elastic wave equation is cast as a generalized Hermitian eigenproblem, enabling the direct application of the abstract perturbation theory developed for quantum mechanics.

3.3.1 Wave Equation Operator-Theoretic Formulation in Phononic Waveguides

As outlined in §2.1.1, the elastic wave equation in solids can be formulated as

$$\nabla \cdot \; c : \nabla_s \vec{v} = -\omega^2 \rho \vec{v} \,. \tag{3.3}$$

The wave equation has been formulated in §2.1.2 as a generalized Hermitian eigenproblem in abstract operator form as

$$\hat{A} \left| \psi \right\rangle = \lambda \hat{B} \left| \psi \right\rangle , \tag{3.4}$$

with the inner product defined as

$$\langle \vec{\mu} | \vec{\nu} \rangle = \int_{\Omega} \vec{\mu}^* \cdot \vec{\nu}, \tag{3.5}$$

where the integral is over the entire solution domain Ω, which in this case refers to the entire waveguide volume. This inner product is defined both for the velocity and stress vector fields, and as shown in [70], we have

$$\left\langle \vec{v}\,' \middle| \nabla \cdot \vec{T} \right\rangle = - \left\langle \nabla_s \vec{v}\,' \middle| \vec{T} \right\rangle , \tag{3.6}$$

which directly implies that $(\nabla \cdot\,)^\dagger = (-\nabla_s)$, from which the Hermiticity of the master operator ($\hat{A} = \nabla \cdot \; c : \nabla_s \vec{v}\,$) in (3.3) follows.

Standard perturbation theory can be directly applied to the operators in (3.3) as in [70]. However, for treatment of waveguide perturbation, a more convenient formulation is adopted with explicit analogy to time-dependent perturbation theory in quantum mechanics [69]. Such treatment has been developed earlier for Maxwell's equations for photonic waveguides and gratings by Johnson, et al. [86, 87]. In this work, the treatment is extended to cover phononic waveguides as well.

We start by considering a master wave equation involving both the velocity and stress fields, with separate terms for propagation direction derivatives. Dirac notation is still used for this purpose; however, the abstract state ket $|\psi\rangle$ is now considered to be

$$|\psi\rangle = \begin{bmatrix} \vec{v} \\ \vec{T} \end{bmatrix} , \tag{3.7}$$

which is a 9×1 vector, representing the full velocity and stress field solution at a given transverse plane; in clear contrast to the electromagnetic case in [86, 87]. In the electromagnetic formulation, the fields along the propagation direction can be inferred from the transverse fields due to the form of Maxwell's equations, which involves curl ($\boldsymbol{\nabla}\times$) spatial derivatives. However, the elastic wave equation (3.3) allows for *longitudinal* elastic waves, where fields along the propagation direction cannot be completely determined form the transverse fields. For an intuitive example, one can think of a material with a trivial Poisson's ratio of 0. In this case the stiffness matrix c becomes diagonal and the material supports longitudinal waves with trivial traverse fields components. There is no analog of such longitudinal waves in electromagnetics.

Define the inner product $\langle\psi'|\psi\rangle$ for the aforementioned state ket as

$$\langle\psi'|\psi\rangle = \int_{\partial\Omega} \vec{v}'^{\,*} \cdot \vec{v} + \vec{T}'^{\,*} \cdot \vec{T} \;\; , \tag{3.8}$$

where $\partial\Omega$ is the waveguide surface normal to direction of propagation.

The equations of motion (2.5) in terms of $|\psi\rangle$, become

$$\begin{pmatrix} \rho\,\mathbb{1}_{3\times3} & \frac{i}{\omega}\nabla\cdot \\ \frac{i}{\omega}\nabla_s & s \end{pmatrix} |\psi\rangle = 0, \tag{3.9}$$

where $\mathbb{1}_{3\times3}$ is the 3×3 identity matrix.

Without loss of generality, we assume $\hat{z}$ to be the propagation direction down the waveguide. We can isolate the derivatives with respect to the propagation direction by decomposing the differential operators as

$$\nabla\cdot\ = \nabla_{tr}\cdot\ + \mathbb{z}\,\partial_z \qquad \text{and} \qquad \nabla_s = \nabla_{tr-s} + \mathbb{z}^{\mathrm{T}}\,\partial_z\ , \tag{3.10a}$$

with $\nabla_{tr-s} = (\nabla_{tr}\cdot)^{\mathrm{T}}$,

$$\nabla_{tr}\cdot\ = \begin{pmatrix} \partial_x & 0 & 0 & 0 & 0 & \partial_y \\ 0 & \partial_y & 0 & 0 & 0 & \partial_x \\ 0 & 0 & 0 & \partial_y & \partial_x & 0 \end{pmatrix} \qquad \text{and} \qquad \mathbb{z} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}. \tag{3.10b}$$

Substituting in (3.9) and moving all z-derivatives to the right hand side, we get the master equation

$$\hat{A}\,|\psi\rangle = -i\,\frac{\partial}{\partial z}\hat{Z}\ |\psi\rangle \tag{3.11a}$$

with

$$\hat{A} = \begin{pmatrix} \omega\,\rho\,\mathbb{1}_{3\times3} & i\,\nabla_{tr}\cdot \\ i\,\nabla_{tr-s} & \omega\,s \end{pmatrix} \quad \text{and} \quad \hat{Z} = \begin{pmatrix} \mathbb{0}_{3\times3} & \mathbb{z} \\ \mathbb{z}^{\mathrm{T}} & \mathbb{0}_{6\times6} \end{pmatrix}, \tag{3.11b}$$

where $\mathbb{0}_{n\times n}$ is the $n\times n$ all-zeros matrix. $\hat{A}$ and $\hat{Z}$ are both Hermitian under the inner product (3.8) for lossless materials (with real s). Hermiticity proof of $\hat{A}$ is provided in Appendix A. It is important to note that no approximation has been made to derive (3.11), aside from assuming the materials involved are linear elastic lossless materials. This formulation highlights the linear operator nature of the elastic waves while abstracting the detailed underlying derivatives and tensor manipulations.

It is straightforward to prove that

$$\begin{aligned}\frac{1}{2\omega}\langle\psi|\,\hat{A}\,|\psi\rangle &= \frac{1}{2}\int_{\partial\Omega}\rho\,\vec{v}^{\,*}\cdot\,\vec{v}\;\;+\;\vec{T}^{\,*}\!: s : \vec{T}\\ &= u_s + u_v\\ &= U,\end{aligned} \tag{3.12}$$

where u_s and u_v are the strain and kinetic energy density in the waveguide per unit length and U is the total stored energy per unit length [68]. One also finds that

$$\begin{aligned}\frac{1}{4}\langle\psi|\,\hat{Z}\,|\psi\rangle &= \frac{1}{4}\int_{\partial\Omega}\vec{v}^{\,*}\hat{Z}\vec{T}\;+\vec{T}^{\,*}\hat{Z}\vec{v}\\ &= \frac{1}{4}\int_{\partial\Omega}(v_x^*T_{xz}+v_y^*T_{yz}+v_z^*T_{zz})+(T_{zz}^*v_z+T_{yz}^*v_y+T_{xz}^*v_x)\\ &= \int_{\partial\Omega}\vec{\mathcal{P}}\,\cdot\,\hat{z}\\ &= P,\end{aligned} \tag{3.13}$$

where $\vec{\mathcal{P}}$ is the Poynting vector and P is the average power travelling down the waveguide [68]. The choice of the state ket $|\psi\rangle$ together with the inner product definition results in an operator-theoretic framework that is closely related to energy and power, as well as the general physics of the elastic waves.

The major advantage of the formulation in (3.11) becomes apparent when considering its explicit similarity to the time-dependent Schrödinger equation, with z taking the place of time. The general properties of Hermitian eigenproblems are directly valid for (3.11), without explicit derivations involving cumbersome tensor algebra. Moreover, the results of well-established time-dependent perturbation theory in quantum mechanics immediately apply to (3.11), with z playing the role of the time variable. This allows us to easily study the effects of perturbation along the waveguide propagation direction.

3.3.2 Coupled Mode Theory for Non-Grated Phononic Waveguides

Coupled mode theory provides a formal solution for (3.11) that allows us to investigate the effect of different perturbations on the waveguide along the propagation direction. It has been previously investigated both for photonics and elastic waves [86–89]. The presented framework has *not* previously been considered for elastic waves.

Uniform, Non-grated Waveguides

Consider a hypothetical infinitely long waveguide with a uniform cross-section along $\hat{z}$. The solution to (3.11) can be assumed to take the form $e^{i\beta z}\left|\psi_{\mathrm{tr}}\right\rangle$, where $\left|\psi_{\mathrm{tr}}\right\rangle$ is a function only of the cross-section coordinates. For uniform waveguides, (3.11) reduces to

$$\hat{A}\left|\psi_{\mathrm{tr}}\right\rangle = \beta\hat{Z}\left|\psi_{\mathrm{tr}}\right\rangle . \tag{3.14}$$

It is important to note that $\hat{Z}$ is Hermitian, yet it is not positive definite. The orthogonality relation in this case is given by

$$\left\langle\beta^*\right|\hat{Z}\left|\beta'\right\rangle = 0, \quad \forall\, \beta \neq \beta', \tag{3.15}$$

where $\left\langle\beta^*\right|$ is the *bra* corresponding to $\left|\beta^*\right\rangle$, the state with the conjugated eigenvalue β^* [86]. The conjugation is necessary since $\hat{Z}$ is not positive definite, and hence β is allowed to be complex. Modes with imaginary β values correspond to evanescent modes. The eigenmodes are normalized such that

$$\left\langle\beta^*\right|\hat{Z}\left|\beta'\right\rangle = \delta_{\beta\beta'}. \tag{3.16}$$

An additional phase might be added to such normalization.

Instantaneous Eigenmodes

Assume that small (and slow) perturbations are introduced along the length of the waveguide. For sufficiently slow perturbation, at each z, one can consider a *virtual* uniform and infinite waveguide that has the same cross section $\partial\Omega_z$ as the original one. The eigenmodes of this virtual waveguide are solutions to (3.14) over the cross section $\partial\Omega_z$, forming a complete and orthogonal basis. Focusing only on the guided modes, the eigenmodes form a discrete spectrum which we can label $|n; z\rangle$, where n is a mode index and the subscript z refers to the location z along the original waveguide. Adopting this notation, the master equation (3.14) can be written for every z as

$$\hat{A}(z)\,|n; z\rangle = \beta_n(z)\hat{Z}\,|n; z\rangle\,. \tag{3.17}$$

The eigenmodes $|n; z\rangle$ can be thought of as *instantaneous* (at a given z, rather than at a given time) eigenmodes of the original waveguide and can be used as a basis for expanding the solution of the full waveguide, with z-dependent expansion coefficients $c(z)$. The perturbed waveguide solution can be assumed to take the form

$$|\psi(z)\rangle = \sum_n c_n(z)e^{i\theta_n(z)}\,|n; z\rangle\,, \tag{3.18}$$

where

$$\theta_n(z) = \int_z dz'\,\beta_n(z'). \tag{3.19}$$

This phase is selected in the above integral form to match the *Berry phase* from quantum mechanics [69]. Its derivative in the z-direction is given by

$$\frac{d\,\theta_n(z)}{d\,z} = \beta_n(z). \tag{3.20}$$

Expansion Coefficients in $|n; z\rangle$ Basis

Substituting this full solution into (3.14) yields

$$\begin{aligned} -i\frac{\partial}{\partial z}\hat{Z}\,|\psi(z)\rangle &= \hat{Z}\sum_n e^{i\theta_n(z)}\left(-i\frac{dc_n(z)}{dz}\,|n;z\rangle - i\,c_n(z)\frac{\partial\,|n;z\rangle}{\partial z} + \beta_n(z)c_n(z)\,|n;z\rangle\right) \\ &= \hat{A}\,|n;z\rangle \\ &= \hat{Z}\sum_n \beta_n(z)c_n(z)e^{i\theta_n(z)}\,|n;z\rangle\,. \end{aligned} \tag{3.21}$$

This simplifies to

$$\hat{Z}\sum_n e^{i\theta_n(z)}\left[\frac{dc_n(z)}{d\,z}\,|n;z\rangle + c_n(z)\frac{\partial}{\partial_z}\,|n;z\rangle\right] = 0. \tag{3.22}$$

Taking the inner product with $\langle m^*; z|\, e^{-i\theta_m(z)}$ and using the orthogonality of the eigenmodes of the waveguide at equal z, we get

$$\frac{d\,c_m(z)}{d\,z} = -\sum_n c_n(z)e^{i(\theta_n(z)-\theta_m(z))}\,\langle m^*;z|\,\hat{Z}\frac{\partial}{\partial_z}\,|n;z\rangle\,. \tag{3.23}$$

This is a system of differential equations that can be solved in space for the expansion coefficients $c_n(z)$. The right-hand-side of (3.23) can be further expanded. First, consider the derivative of (3.17) with respect to z,

$$\frac{\partial\hat{A}}{\partial_z}\,|n;z\rangle + \hat{A}\frac{\partial}{\partial_z}\,|n;z\rangle = \frac{\partial\beta_n(z)}{\partial_z}\hat{Z}\,|n;z\rangle + \beta_n(z)\hat{Z}\frac{\partial}{\partial_z}\,|n;z\rangle\,. \tag{3.24}$$

Taking the inner product with $\langle m^*; z|$ for the case $m \neq n$, equation (3.24) reduces to

$$\langle m^*;z|\,\frac{\partial\hat{A}}{\partial_z}\,|n;z\rangle = \big(\beta_n(z) - \beta_m(z)\big)\,\langle m^*;z|\,\hat{Z}\frac{\partial}{\partial_z}\,|n;z\rangle\,. \tag{3.25}$$

Substituting back into (3.23), we get

$$\frac{d\,c_m(z)}{d\,z} = -c_m(z)\,\langle m^*;z|\,\frac{\partial}{\partial_z}\,|m;z\rangle - \sum_n c_n(z)e^{i(\theta_n(z)-\theta_m(z))}\frac{\langle m^*;z|\,\frac{\partial\hat{A}}{\partial_z}\,|n;z\rangle}{\beta_n(z)-\beta_m(z)}. \tag{3.26}$$

The inner product in the second term can be further expanded as

$$\langle m^*; z|\, \frac{\partial \hat{A}}{\partial_z}\, |n; z\rangle = \omega \int_{\partial\Omega} \frac{\partial \rho}{\partial_z}\, \vec{v}^{\,*}_{m^*} \cdot \vec{v}_n + \vec{T}^{\,*}_{m^*} : \frac{\partial s}{\partial_z} : \vec{T}_n$$
$$+ \frac{i}{\omega}\, \vec{v}^{\,*}_{m^*} \cdot \nabla_{tr} \cdot \vec{T}_n + \frac{i}{\omega}\, \vec{T}^{\,*}_{m^*} \cdot \nabla_{tr-s} \vec{v}_n \qquad (3.27)$$

The solution (3.18) with the expansion coefficients (3.26) is still a formal solution for (3.17) in the case of *slow* perturbations and is known as the *coupled mode theorem*. It actually demonstrates that, as the wave propagates down the waveguide, modes $|n; z\rangle$ with $n \neq m$ will start to couple to $|m; z\rangle$ due to the second term. This coupling is basically the *scattering* that arises in phononic waveguides due to perturbations therein.

3.3.3 Coupled Mode Theory for Grated Waveguides

Grated waveguides, such as those constructed from PnCs, demonstrate similar characteristics to those of uniform waveguides. Coupled mode theory in grated waveguides gives qualitatively similar results as (3.26) in addition to more interesting details about periodicity. Derivation of coupled mode theory is included after [86] with highlights on its implications for resonators.

Master Equation in Grated Waveguides

In the case of a periodic waveguide with period Λ, Bloch's theorem suggests a solution of the form $e^{i\beta z}\, |\beta\rangle$, where $|\beta\rangle$ is a periodic function of z with period a. The master equation for this periodic waveguide reduces to

$$\hat{A}\, |\beta\rangle = \beta \hat{Z}\, |\beta\rangle - i \frac{\partial}{\partial z}\, |\beta\rangle$$
$$\hat{C}\, |\beta\rangle = (\hat{A} + i \frac{\partial}{\partial z} \hat{Z})\, |\beta\rangle = \beta \hat{Z}\, |\beta\rangle\,. \qquad (3.28)$$

From Bloch's theorem, solutions with eigenvalues $\beta + 2\pi l/\Lambda,\ \forall l \in \mathbb{N}$ are equivalent. It is thus sufficient to consider only the first Brillouin zone with $\beta \in [-\pi/\Lambda, \pi/\Lambda)$.

This also results in the equivalence relation

$$e^{i\beta z}\left|\beta\right\rangle = e^{i\beta z}e^{i2\pi\, lz/\Lambda}\left|\beta+\frac{2\pi\, l}{\Lambda}\right\rangle \quad \text{or} \quad \left|\beta+\frac{2\pi\, l}{\Lambda}\right\rangle = e^{-i2\pi\, lz/\Lambda}\left|\beta\right\rangle . \tag{3.29}$$

Instantaneous Basis for Perturbed Grated Waveguides

As we did in §3.3.2, we will assume small and slow perturbation along the waveguide propagation direction z. Such perturbation is not only limited to material properties, but it can also be a period perturbation; hence we will denote the period $\Lambda(z)$. Again, we associate a hypothetical, infinitely extending, unperturbed, grated waveguide that coincides with the physical structure at each position z. We will use $\bar{z}$ to denote the coordinates of the unperturbed waveguides. This is necessary since the uniform waveguide is itself grated. We also define a normalized coordinate $\bar{w} = \bar{z}/\Lambda(z)$, so that all the virtual waveguides have a unit period in $\bar{w}$. The corresponding coordinate in the physical waveguide will be denoted w and is defined as

$$w = \int^{z} \frac{dz'}{\Lambda(z')}, \tag{3.30}$$

which corresponds to counting the periods up to z in the physical waveguide.

The virtual waveguide is selected to coincide with the physical waveguide at $\bar{w} = w$ which corresponds to

$$\bar{s}(z, w) = s(z) \quad \text{and} \quad \bar{\rho}(z, w) = \rho(z), \tag{3.31}$$

where $\bar{s}(z, \bar{w})$ and $\bar{\rho}(z, \bar{w})$ are the compliance matrix and material density of the virtual waveguide, parameterized by z. This is equivalent to specifying an origin in the virtual waveguide space. The instantaneous virtual waveguide has eigenmodes $\left|n(z, \bar{w})\right\rangle$ that satisfy

$$\hat{C}(z, \bar{w})\left|n(z, \bar{w})\right\rangle = \left(\hat{A}(z, \bar{w}) + \frac{i}{\Lambda(z)}\frac{\partial}{\partial \bar{w}}\hat{Z}\right)\left|n(z, \bar{w})\right\rangle = \beta_n(z)\hat{Z}\left|n(z, \bar{w})\right\rangle , \tag{3.32}$$

where $d\bar{z} = \Lambda(z)d\bar{w}$ has been used*.

Now, the solution of the physical waveguide $|\psi(z)\rangle$ should be expanded in terms of the eigenmodes of the virtual waveguides. It is instructive to expand in terms of $|n(z,w)\rangle$, i.e. at $\bar{w} = w$, where the instantaneous virtual waveguide coincides with the physical structure. However, this removes the explicit dependence on the coordinates $\bar{w}$ (as we set $\bar{w} = w$) which is necessary for the orthogonality relation resulting from (3.32), where the integration should be carried over $\bar{w}$. The solution for this problem proposed in [86] is to use a shifted basis in the virtual waveguide space $|n(z, w + \bar{w})\rangle$, which solves (3.32) for the operator $\hat{A}(z, w+\bar{w})$. Treating the shift $\bar{w}$ as a parameter, and considering $\hat{A}(z, w + \bar{w})$ as a function of z, we get a family of physical problems. The solution to these physical systems $|\psi(z, \bar{w})\rangle$ exactly coincides with the solution for the physical system at $\bar{w} = 0$. Thus, we can solve for the evolution of $|\psi(z, \bar{w})\rangle$ over z with $\bar{w}$ being a parameter. (Compare this to the simpler case of §3.3.2.) We can thus expand $|\psi(z, \bar{w})\rangle$ as

$$|\psi(z,\bar{w})\rangle = \sum_n c_n(z,\bar{w})\, |n(z, w+\bar{w})\rangle\, e^{i\theta_n(z)}, \tag{3.33a}$$

with

$$\theta_n(z) = \int^z dz'\, \beta_n(z'). \tag{3.33b}$$

Owing to the unit periodicity of the virtual waveguides in $\bar{w}$, the coefficients $c_n(z, \bar{w})$ are also periodic with unit period and can be expanded by means of Fourier series as

$$c_n(z,\bar{w}) = \sum_l c_{n,l}(z) e^{-i\,2\pi l\bar{w}}, \tag{3.34}$$

with $c_{n,l}(z)$ being the Fourier coefficients and the physical solution corresponds to $c_n(z, 0) = \sum_l c_{n,l}$.

*$\Lambda(z)$ is a fixed parameter for the virtual waveguide and hence it does not contribute to the derivatives in this equation.

Coupled Mode Coefficients

To obtain the defining equations for $c_{n,l}$, we substitute the solution (3.33a) into the master equation (3.11) as we did in §3.3.2. The expansion after [86] yields

$$\begin{aligned} -i\frac{\partial}{\partial z}\hat{Z}\,|\psi\rangle &= \hat{Z}\sum_n e^{i\theta_n}\left[-i\frac{dc_n}{dz}\,|n\rangle - ic_n\frac{\partial\,|n\rangle}{\partial z} - i\frac{c_n}{\Lambda}\frac{\partial\,|n\rangle}{\partial\bar{w}} + \beta_n c_n\,|n\rangle\right] \\ &= \hat{A}(z, w+\bar{w})\,|\psi\rangle \\ &= \hat{Z}\sum_n\left[-i\frac{c_n}{\Lambda}\frac{\partial\,|n\rangle}{\bar{w}} + \beta_n c_n\,|n\rangle\right], \end{aligned} \tag{3.35}$$

where* $|\psi\rangle \equiv |\psi(z,\bar{w})\rangle$, $c_n \equiv c_n(z,\bar{w})$ and $|n\rangle \equiv |n(z, w+\bar{w})\rangle$. Taking the inner product with $\langle m^*(z, w+\bar{w})|\, e^{-i\theta_m} e^{i\,2\pi k\bar{w}}$, we can separate a system of differential equations in terms of the k^{th} Fourier expansion coefficient for the m^{th} mode $c_{m,k}$. The coupled mode equations become

$$\frac{dc_{m,k}}{dz} = -\sum_{n,l} c_{n,l}e^{i(\theta_n-\theta_m)}\,\langle m*|\,\hat{B}e^{i\,2\pi(k-l)\bar{w}}\frac{\partial}{\partial z}\,|n\rangle \tag{3.36}$$

Finally, as described in [86], the inner product on the right-hand-side is found in terms of the derivative of the master operator $\hat{C}$,

$$\frac{dc_{m,k}}{dz} = -c_{m,k}\,\langle m^*|\,\hat{Z}\frac{\partial}{\partial z}\,|m(z)\rangle - \sum_{n\neq m, l\neq k}\frac{\langle m^*|\,e^{i\,2\pi k\bar{w}}\partial_z\hat{C}_z(\bar{w})e^{-i2\pi l\bar{w}}\,|n\rangle}{\Delta\beta_{n,l;m,k}}, \tag{3.37}$$

where the difference in wave number $\Delta\beta_{n,l;m,k}$ is given by

$$\Delta\beta_{n,l;m,k}(z) = \beta_n(z) - \beta_m(z) + \frac{2\pi}{\Lambda}(l-k), \tag{3.38}$$

and $\partial_z\hat{C}_z$ is

$$\frac{\partial\hat{C}}{\partial z} = \frac{\partial\hat{A}}{\partial z} + i\frac{d\Lambda^{-1}}{dz}\hat{Z}\frac{\partial}{\partial\bar{w}}. \tag{3.39}$$

It is important to note that

---

* The partial derivative w. r. t. z on the first line is considered a total z derivative when applied to $|n(z, w+\bar{w})\rangle$; hence the expansion $\frac{d}{dz}\,|n(z, w+\bar{w})\rangle \equiv \partial_z\,|n(z, w+\bar{w})\rangle + (i/\Lambda)\partial_{\bar{w}}\,|n(z, w+\bar{w})\rangle$.

- The commutator bracket of the exponential phase term and $\bar{w}$ derivative is a constant that gets integrated to zero in the inner product

$$\left[\frac{\partial}{\partial \bar{w}}, e^{-2\pi\, i\, l\, \bar{w}}\right] = -2\pi\, i\, l e^{-2\pi\, i l \bar{w}}. \tag{3.40}$$

  This allows moving the phase term about the derivative of $\hat{C}$.

- In the limit of $\Lambda \to \infty$, we recover the results of the non-grated waveguides of §3.3.2.

- Equation (3.39) is similar to (3.27) with the additional term to account for waveguide period changes.

3.4 Scattering in Phononic Waveguides

As discussed in §3.3.2 and §3.3.3, perturbations along the length of the transmission line result in scattering, which is basically the coupling of the different waveguide modes. Such scattering is highly undesirable in MEMS resonators. Coupling to non-guided modes (radiation modes) such as plane waves in the bulk or PnC modes results in radiation energy losses from the resonator. A degradation of the resonator quality factor is the most observable effect on the resonator performance. Furthermore, scattering may lead to coupling into other guided modes, which might result in resonances at different frequencies. This is directly manifested as undesirable spurious modes in the resonator frequency response.

For this reason, it is highly desirable to reduce the scattering in the phononic waveguide. Perturbations resulting from random process variations can be minimized by using a smaller resonator area for better local process uniformity. On the other hand, systematic perturbation may arise by virtue of the resonator design. Abrupt truncation of the waveguide structure, terminations and non-uniform structures are examples of systematic perturbations. It is thus important to understand the characteristics and sources of scattering.

The main characteristics of scattering are found from the expansion coefficients in (3.26). Two major factors can be observed: the difference in the wave vector $\vec{k}$ and the spatial rate of change of the perturbation as discussed below.

3.4.1 $\vec{k}$-Difference and Fully-Differential Operation

The second term in (3.26) represents the mixing between the different waveguide modes. Scattering in phononic waveguides increases with the magnitude of the second term in (3.26). The first dependence that can be observed is the inverse dependence on $\beta_m(z) - \beta_n(z)$, which are the wave vector components along the propagation direction in the waveguide. Scattering is inversely proportional to the distance between the corresponding modes in the $\vec{k}$-space. Based on this inverse dependence, it is highly desirable to increase the distance between the main resonator mode and the undesired radiation/spurious modes in $\vec{k}$-space.

By examining the waveguide dispersion relation in Figure 3-1 on page 82, the shear sound line ($\omega = c_{\text{shear}}\beta$) of the bulk wafer is the closest in $\vec{k}$-space to the guided mode at any given frequency that is well within the PnC bandgap. Thus, selecting the resonance mode to be the furthest possible from the sound line is necessary. This can be achieved by enforcing the wave guide to operate with* $\beta = \pi/a$.

Physically, operating at $\beta = \pi/a$ corresponds to the fields in neighboring periods of the waveguide to be 180° out-of-phase. As will be seen in the next chapter, this can be enforced by using fully differential driving for the resonator.

It is also interesting to note that operating at $\beta = \pi/a$ results in the largest imaginary wave vector k_{tr} in the bulk wafer. This can be seen from the dispersion relation in Figure 3-1 on page 82 as

$$k_{\text{tr}} = k_y = i\sqrt{\left(\frac{\pi}{a}\right)^2 - \left(\frac{\omega}{c}\right)^2}. \tag{3.41}$$

The largest imaginary k_y corresponds to the fastest evanescent wave decay in the bulk substrate. This directly relates to better confinement and higher stress fields at the

*The propagation direction for the waveguide of Figure 3-1 on page 82 is $\hat{x}$ leading to $\beta = k_x$.

MOS transistor gates. Based on these advantages, resonators with fully-differential driving and $\beta = \pi/a$ are more preferable for maximizing Q and minimizing spurious modes.

3.4.2 Spatial Rate of Change of $\hat{A}$ and $\hat{C}$

Another important factor in determining scattering in phononic waveguides is the inner product $\langle m^*; z| \frac{\partial \hat{A}}{\partial_z} |n; z\rangle$ as seen from (3.26). Thus, the rate of change of the master operator $\hat{A}$ has a major impact on the scattering performance of the phononic waveguide. As demonstrated in (3.27), this rate of change is directly proportional to the rate of change of the material properties $\frac{\partial \rho}{\partial_z}$ and $\frac{\partial s}{\partial_z}$ along with a mode overlap integral in non-grated waveguides. It is also interesting to note the rather similar dependence on the waveguide period for grated waveguide as obtained in (3.39).

This implies that abrupt and fast changes in the waveguide cross section, material properties or periodicity will produce more scattering than smooth and slow transitions. Abrupt terminations, discontinuities and sudden coupling to other waveguides are all examples of abrupt spatial changes that should be minimized in order to reduce scattering. Slow transitions and coupling should be used instead. This is a very important result that leads directly to the adiabatic theorem in §3.5.

For the resonators in §3.2, the structure can also be considered as a waveguide along the non-resonant dimension (z-direction in Figure 3-3). Abrupt transitions and perturbations should also be minimized along this non-resonant dimension. This effect was studied by comparing two identical PnC-RBT as described in §3.2. The two RBTs differ in the geometry of the MOSCAPs and sensing transistor source and drain contacts. One of the RBTs uses rectangular wall-like contacts, whereas the other uses standard square CMOS contacts. Measurement results of the two devices are compared in Figure 3-9. The two resonators are believed to be closely matched in terms of the layer thicknesses and dimensions as evident from the perfect alignment of their resonance frequencies.

The resonator with wall-like rectangular vias and a uniform out-of-plane structure shows a higher quality factor and fewer spurious modes than the square vias resonator.

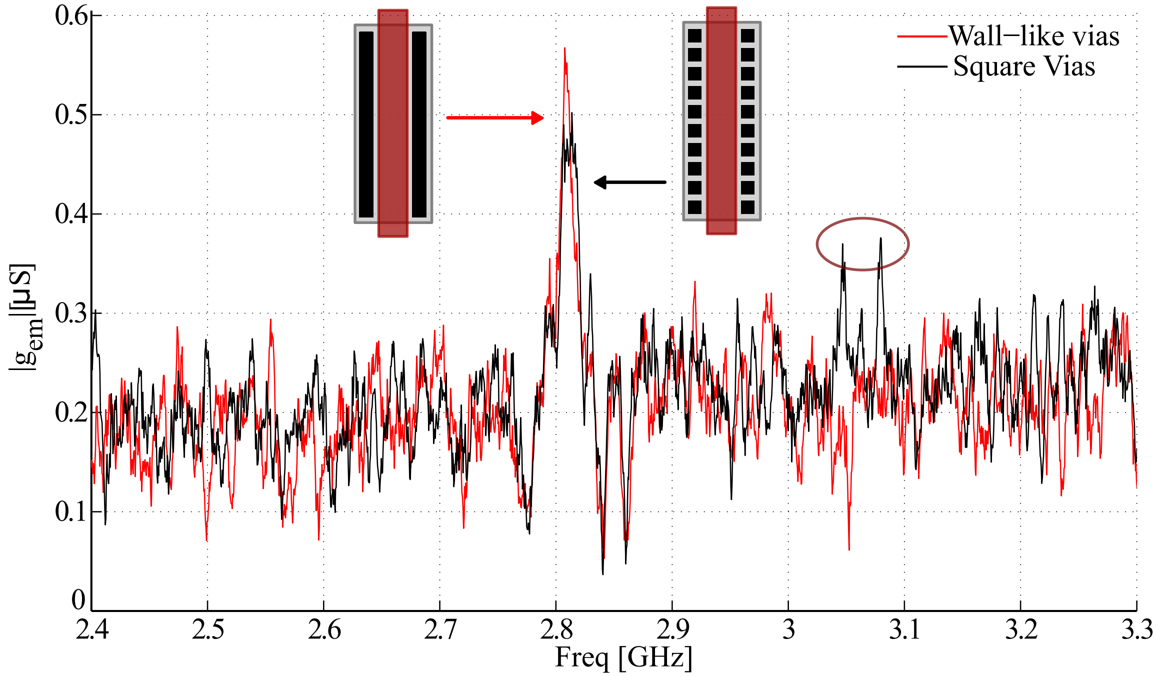


Figure 3-9: Measured electromechanical transconductance g_{em} for identical RBTs with rectangular (wall-like) vias and the regular CMOS square vias. The device with square vias has clearly reduced quality factor Q and extra spurious modes.

The scattering dependence on abrupt variations can be used to qualitatively explain this result. Square vias represent abrupt changes as opposed to the more uniform structure of the wall vias, resulting in more scattering. Ideally, slow terminations should be used in the non-resonant dimensions as well; however, the CMOS technology doesn't allow for such irregular geometries to be implemented.

It is also interesting to note that the drive and sense transducers of the acoustic bragg reflectors RBTs in [36] each spanned only half the cavity depth (in z-direction). This can also be considered a source of significant scattering which contributed to the radiation losses and reduced Q. The uniform cross-section of the RBTs considered in this work is clearly a major advantage over [36].

3.5 The Adiabatic Theorem

Multiple versions of the adiabatic theorem (or adiabatic approximation) have been derived for quantum mechanics as well as photonic waveguides [69,86,90–92]. A proof of the adiabatic theorem for phononic waveguides is included here based on the work of Johnson, *et. al.* [86] for photonic waveguides.

3.5.1 The Adiabatic Theorem

The adiabatic theorem is the assertion that scattering practically vanishes in the limit of sufficiently slow perturbation. In quantum mechanics, this is usually stated as the system remaining in the instantaneous eigenstate without undergoing a transition for slow enough time-dependent perturbations. In photonic waveguides, it basically refers to the propagation down the waveguide, being constrained to the same initial mode, which slowly evolves with the structure to match the local eigenmode. No scattering to other modes occur for adiabatic transitions. The same concept is applicable for the phononic waveguides considered here. Formally, this is expressed in terms of the expansion coefficients of the different eigenstates $c(z)$ as

$$c(z) \to c(0) \quad \text{as} \quad L \to \infty, \tag{3.42}$$

where L is the physical length of the perturbation or transition region. Traditionally a condition is usually imposed for the adiabatic theorem to hold: a *gap* should exist between the different eigenstates, limiting the adiabatic theorem to discrete non-degenerate spectra. However, this restriction has been revisited and a generalization has been demonstrated in [92].

The adiabatic theorem proof follows directly from (3.26) and (3.37). The coupled mode theorem takes the general form of

$$\frac{dc_m}{dz} = \sum_{n \neq m} c_n \left\langle m^* \left| \frac{\partial \hat{A}}{\partial z} \right| n \right\rangle \exp\left(i \int^z \Delta\beta_{nm}(z') d\, z' \right), \tag{3.43}$$

where the intrinsic inner product $\langle m^* | \, \partial_z \hat{A} \, | m \rangle$ can be set to zero by proper phase selection as shown in Appendix B of [86]. By applying a coordinate transform of the form $w = z/L$, equation (3.43) becomes

$$\frac{dc_m}{dw} = \sum_{n \neq m} c_n \left\langle m^* \middle| \frac{\partial \hat{A}}{\partial w} \middle| n \right\rangle \exp\left(iL \int^w \Delta\beta_{nm}(w')d\,w' \right). \tag{3.44}$$

The only dependence on the physical length appears in the exponential phase term. In the limit as $L \to \infty$, with real β (corresponding to guided non-evanescent modes), the exponential phase term oscillates infinitely rapidly and averages to 0, yielding

$$\lim_{L \to \infty} \frac{dc_m}{dw} = 0 \quad \Rightarrow \quad c_m(z) \to c_m(0), \tag{3.45}$$

which is the formal statement of the adiabatic theorem.

3.5.2 Adiabatic Transitions and Tapers

Based on the adiabatic theorem, it becomes evident that the key to reduced scattering is slow transitions. Adiabatic tapers have been adopted in many photonic waveguides designs to achieve efficient bending and coupling between different waveguides [71, 86, 87, 93–97]. Reducing scattering in these coupling setups ensures better insertion loss (by lowering scattering to radiation modes) as well as reduced dispersion (by maintaining single-mode operation). The adiabatic theorem has been concerned with guided propagating modes (with real β). As discussed in [86], if the initial mode falls into the bandgap of the instantaneous waveguide along the taper, exponentially decaying evanescent waves result along with reflections.

An adiabatic taper can be intentionally designed so that the mode is in an instantaneous bandgap, creating a reflector instead of a matching structure. Such reflectors have been used in resonant cavity designs, both in photonics and acoustics [98–100]. The advantage of adiabatic tapering in the reflector is reducing scattering, allowing only for specular reflection. This corresponds to coupling to the mode with $\beta_2 = -\beta_1$, with minimal coupling to other undesired modes. As highlighted in §3.4, this reduced scattering translates to a higher quality factor as well as reduced spurious modes in the resonator.

3.6 Conclusion

The implementation of phononic waveguides in CMOS via the use of BEOL phononic crystals has been demonstrated. Waveguiding is achieved by virtue of the wave being unable to propagate in the PnC bandgap and the higher sound velocity in the bulk wafer. This phenomenon is a dual of the index guiding of photonics. Although this study considered IBM 32 nm SOI technology, the SOI buried oxide (BOX) layer plays no role in elastic waveguiding, making this technique directly applicable to bulk technologies. Moreover, the dispersion characteristics of CMOS phononic waveguides can also be engineered through the horizontal dimensions that can be controlled by the designer in most CMOS technologies.

The first unreleased CMOS resonant body transistor based on PnC waveguide has been demonstrated. It achieves a quality factor of 252 at 2.81 GHz, marking an 8× improvement in Q over previous RBTs. A single sensing FET was used along with two MOSCAPs featuring single-ended driving.

Next, coupled mode theory for both non-grated and grated waveguides has been explored. With the goal of reducing scattering in CMOS unreleased resonators, two approaches were suggested:

- Using fully differential driving to operate with $\beta = \pi/a$, the furthest possible from the sound cone $\omega \geq c\,\vec{k}$, and hence reducing the scattering to bulk radiation modes.

- Incorporating slowly varying transitions, even for the reflectors at the waveguide ends, to allow only for specular reflection and to reduce scattering.

The advantage of reducing scattering in the resonator are an increased quality factor and the reduction of spurious modes.

The next chapter explores the implementation and optimization of full RBT devices applying the aforementioned concepts.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 4

CMOS PnC-Based RBTs: Design and Numerical Optimization

Chapter 2 discussed the theory and implementation of phononic crystals in standard CMOS technologies, allowing the realization of good acoustic reflectors therein. Using such PnCs, it has been demonstrated in chapter 3 that it is possible to construct acoustic waveguides capable of confining elastic vibrations in the standard CMOS FEOL layers. In this chapter, CMOS phononic waveguides are used to implement elastic resonant cavities. The concepts of dispersion characteristics engineering and scattering mitigation are used extensively to achieve perfect acoustic confinement, both horizontally and vertically. This chapter demonstrates the general structure of PnC-based CMOS RBTs along with numerical optimization techniques to achieve the best designs. Measurement results for different RBT implementations in IBM 32 nm SOI technology are also reported.

4.1 PnC RBT Resonance Cavity

This section develops the major characteristics and features of the RBT mechanical resonance cavity. The general form of the RBT cavities considered in this work is shown in Figure 4-1. The cavity is formed from multiple phononic waveguide sections: a main cavity as well as termination on both sides.

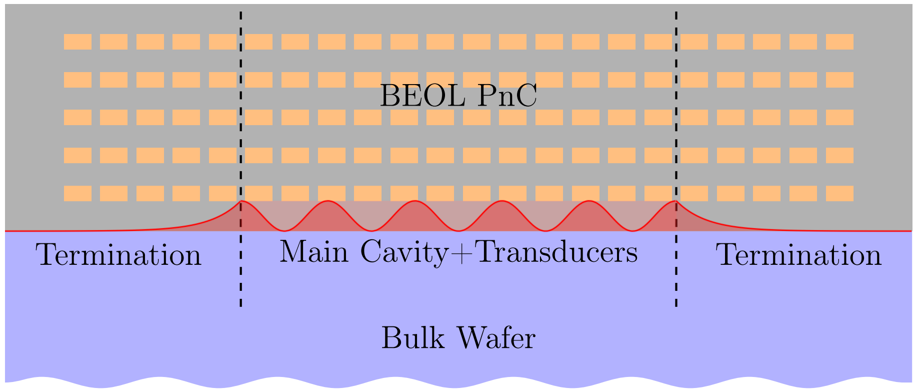


Figure 4-1: General form of the CMOS RBT resonance cavity formed from BEOL PnC phononic waveguides.

4.1.1 Main Cavity Waveguide

As shown in Figure 4-1, the main cavity is the central part of the RBT. It is formed from a phononic waveguide section that has a propagating mode exactly at the operating frequency. On both sides, waveguide termination sections are included to provide the acoustic wave reflection needed for the RBT. A good RBT design will have the highest stresses as well as stored energy confined to the main cavity section and decay exponentially in the termination waveguide sections on both sides. The main cavity also includes the driving and sensing transducers, which for the RBTs considered here are MOSCAPs and MOSFETs, respectively.

While the inclusion of the MOSFETs in the FEOL layers will affect the specific dispersion relation of the waveguide, they do *not* affect the waveguiding concepts presented in §3.1 There, waveguiding was a result of the PnC bandgap preventing the modes from propagating upwards as well as the higher sound velocity in the bulk wafer resulting in index guiding. The dispersion relation for a typical phononic waveguide employed as the main cavity of an RBT is shown in Figure 4-2. The waveguide includes the full MOSFET structure, and the waveguide period was chosen to be double that of the PnC. The dispersion relation still shows the relevant waveguide characteristics as in §3.1.

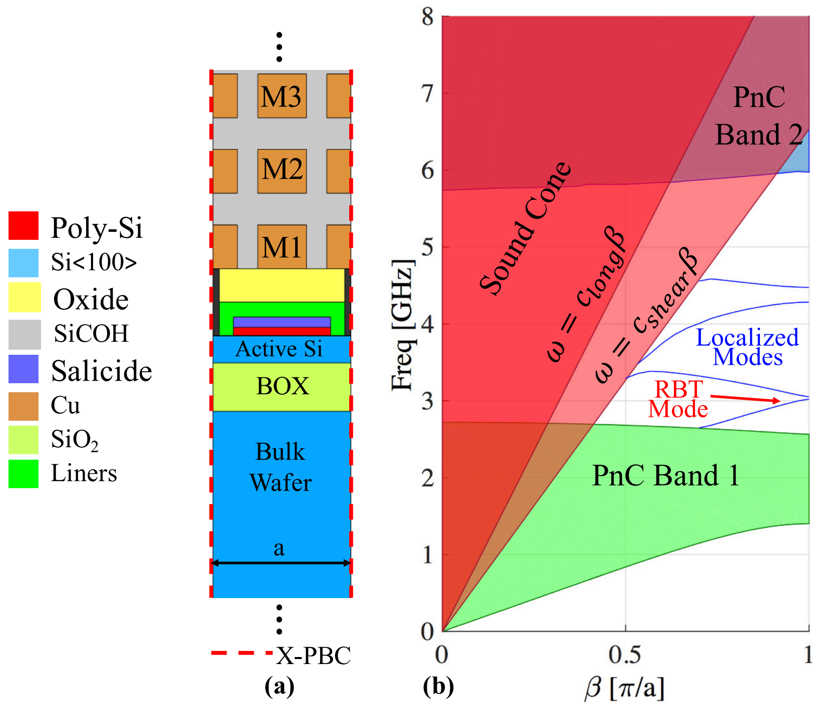


Figure 4-2: Dispersion characteristics for a typical main RBT cavity design, including a complete MOSFET in the FEOL layers. The waveguiding behavior is not affected by the MOSFET structure.

4.1.2 Fully-differential Drive and Sense

As discussed in §3.4, in order to benefit the most from the phononic waveguides, it is important to operate the furthest possible from the sound line, specifically at $\beta = \pi/a$. Operating at a specific β can be enforced by the spatial configuration of the driving transducers (the MOSCAPs). Physically, operating at $\beta = \pi/a$ corresponds to fields in the neighboring waveguide periods being out-of-phase by 180°. This can be easily implemented by fully-differential driving of the resonator, forcing neighboring periods to be out-of-phase.

Ideal In-Phase Single-Ended Driving

To better understand why fully-differential driving generates excitation near $\beta = \pi/a$, we turn to studying the different spatial stress distribution in the reciprocal lattice ($\vec{k}$-space). First consider a perfectly periodic in-phase driving actuating stress distribution like that shown in Figure 4-3.a. The structure period is a, whereas the

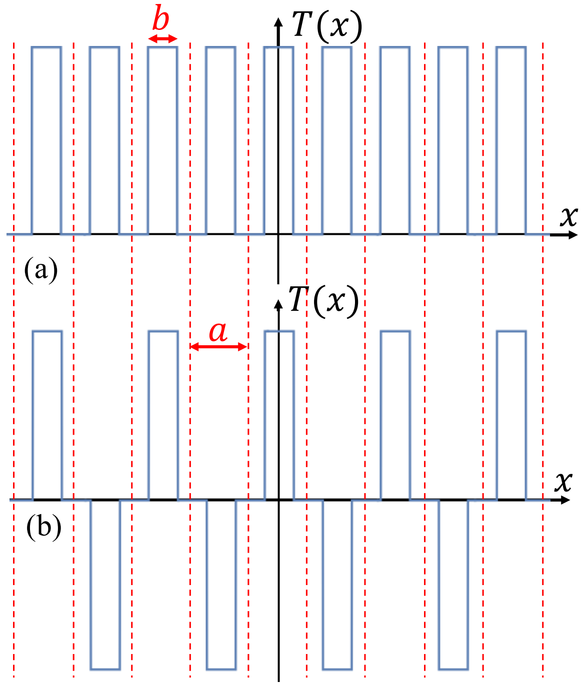


Figure 4-3: Spatial actuation stress distribution for (a) in-phase driving and (b) fully-differential driving.

driving stress is assumed to be uniform over a length $b < a$ (b is approximately the MOSCAP gate length). Such stress distribution can be represented in the real geometrical space by the Heaviside $\Pi(x)$ function as

$$T(x) = T_\circ \, \Pi\left(\frac{x}{b}\right) \ , \ \forall x \in \left[-\frac{a}{2}, \frac{a}{2}\right] \quad \text{and} \quad T(x + n\,a) = T(x) \ , \ \forall n \in \mathbb{Z}. \tag{4.1}$$

The relation between the real and reciprocal lattice is given by Fourier series as

$$T(x) = \sum_{-\infty}^{\infty} T_n \exp\left(i\frac{2\pi n}{a}x\right) \quad ; \quad T_n = \frac{1}{a}\int_{-a/2}^{a/2} dx \exp\left(-i\frac{2\pi n}{a}x\right). \tag{4.2}$$

The Fourier coefficients for in-phase driving stress are thus found to be

$$T_n = T_\circ \frac{\sin(n\pi b/a)}{n\pi} = \frac{T_\circ\, b}{a}\mathrm{sinc}\left(\frac{n\pi b}{a}\right), \tag{4.3}$$

where the coefficients correspond to the $\beta = 2n\pi/a$ in $\vec{k}$-space. The in-phase driving scheme shows the strongest coupling into the sound at $\beta = 0$, with the highest component $T_0 = T_\circ b/a$, or the DC value of the driving stress. This corresponds to the physical picture of radiating uniform plane wave normal to the waveguide propagating downward into the bulk wafer. Moreover, this in-phase coupling has no components near $\beta = \pi/a$ and hence cannot drive the guided modes of the phononic waveguides*.

Ideal Fully-Differential Driving

On the other hand, consider the fully-differential driving scheme of Figure 4-3.b, where neighboring waveguide periods are driven 180° out-of-phase. The driving stress has a spatial period of $2a$, twice that of the waveguide. The stress spatial distribution is given by

$$T(x) = -T_\circ \, \Pi\left(\frac{x+a}{b}\right) \; + \; T_\circ \, \Pi\left(\frac{x}{b}\right) \; - \; T_\circ \, \Pi\left(\frac{x-a}{b}\right) \quad , \quad \forall x \in [-a, a]\,. \tag{4.4}$$

The Fourier series expansion in $\vec{k}$-space becomes

$$T(x) = \sum_{-\infty}^{\infty} T_n \exp\Big(i\frac{n\pi}{a}x\Big) \quad ; \quad T_n = \frac{1}{2a}\int_{-a}^{a} dx\, T(x) \exp\Big(-i\frac{n\pi}{a}x\Big). \tag{4.5}$$

The Fourier series coefficients simplifies to

$$T_n = -i^{3n+1}\, T_\circ \,\mathrm{sinc}\left(\frac{n\pi}{2}\right) \sin\left(\frac{n\,\pi\, b}{2\,a}\right), \tag{4.6}$$

where the coefficients correspond to $\beta = n\pi/a$ in $\vec{k}$-space. The coefficients clearly vanish for even values of $n = \cdots, -2, 0, 2, \cdots$, corresponding to $\beta = 2m\pi/a\ ,\ m \in \mathbb{N}$; leaving only odd n components. The largest component occurs at $n = 1$, corresponding to $\beta = \pi/a$. Thus, fully-differential driving strongly couples to the guided waveguide modes at $\beta = \pi/a$ without coupling into the sound cone†.

There is an intuitive physical explanation as to why this particular driving pat-

*Non-idealities in the structure due to manufacturing may lead to scattering that causes weak coupling into the guided modes.

†With the exception of scattering from manufacturing non-idealities.

tern actually reduces scattering into the bulk wafer sound cone. One can think of a fully-differential transducer array enforcing $\beta = \pi/a$ for the resonator. If an observer is located far away in the bulk wafer, much further than the array size, all the transducers eventually look like point sources. With the observation distance being much larger than the individual separation between the sources, the distance traveled by the waves from all the sources is almost the same. This causes the superimposing waves at the observer to acquire almost the same phase shift, and hence they all arrive 180° out-of-phase with each other. Thus, their superposition vanishes, resulting in no radiation into the bulk wafer.

Fully-Differential Driving in Real Structure

Unlike the ideal structures discussed above, realistic RBTs are not infinitely periodic. The limited structure periodicity strongly affects the expansion coefficients in the $\vec{k}$-space. From the well-known properties of Fourier analysis, the non-periodic structure results in a continuous spectrum in $\vec{k}$-space as opposed to the discrete expansion coefficients of (4.3) and (4.6). We turn to Fourier transform in this case. The driving stress in a finite structure with N periods can be expressed as

$$\tau(x) = T_\circ \Pi\left(\frac{x - a/2}{b}\right) - T_\circ \Pi\left(\frac{x + a/2}{b}\right) \quad , \quad \forall x \in [-a, a]\,; \tag{4.7a}$$

$$T(x) = \sum_{m=0}^{N-1} \tau(x - 2\,m\,a), \tag{4.7b}$$

where $\tau(x)$ has been defined as stress in a fully-differential period. The Fourier transform pair used to relate the real space and the $\vec{k}$-space is defined as

$$\tilde{T}(\beta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx\; T(x) e^{-i\,\beta x} \quad ; \quad T(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\beta\; \tilde{T}(\beta) e^{i\,\beta x}. \tag{4.8}$$

In order to compare the effect of the number of periods, the Fourier decomposition $\tilde{T}(\beta)$ for different number of periods is plotted in Figure 4-4. In order to obtain numerical results, the ratio b/a was assumed to be 75%, which is typical for many RBT designs. As expected, increasing the number driving periods both increases the

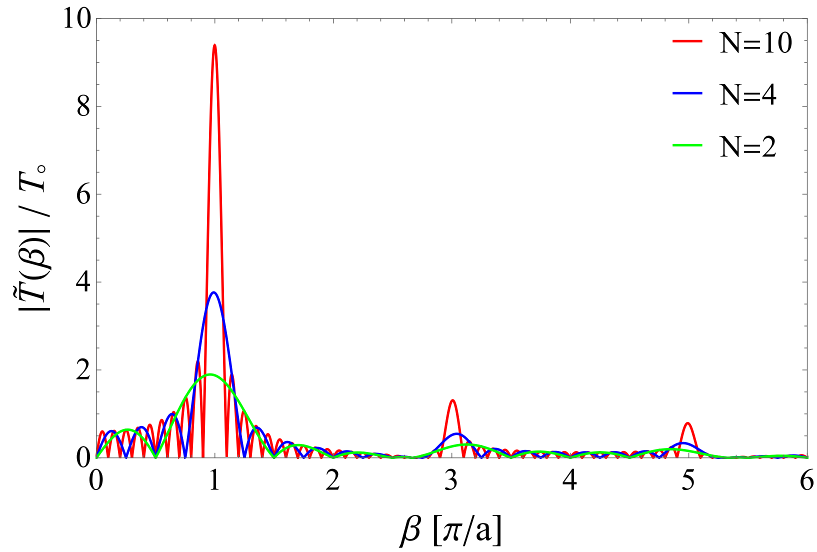


Figure 4-4: Comparison between the $\vec{k}$-space components in case of finite driving transducers, with 2, 5 and 10 periods.

magnitude of the driving component in $\vec{k}$-space as well as resulting in a narrower spectrum. Narrower driving spectrum in $\vec{k}$-space is highly favored to minimize scattering and coupling to the sound cone. However, this comes at the cost of a larger resonator footprint. In most RBTs considered, four fully-differential driving periods were used on each side of the sensing MOSFETs for a total of eight driving periods. This design seems to provide good confinement with reasonable resonator footprint.

Fully-Differential Sensing

As in the case with fully-differential driving, fully-differential sensing is also used, making the resonator a balanced-balanced device. Fully differential sensing also helps in rejecting modes away from $\beta = \pi/a$. The full resonator can be thought of as a three-stage system, namely

1. a driving stage that converts input voltage to stress,
2. mechanical resonance that filters the resulting stress with high selectivity and
3. a sensing stage that converts the stress back to output current.

With both driving and sensing favoring modes close to $\beta = \pi/a$, the effect of scattering and non-idealities is significantly reduced.

4.1.3 Termination Waveguides

The main phononic waveguide structure provides vertical energy confinement for the resonance cavity. Yet it allows waves to propagate freely in the horizontal direction inside the FEOL layers. *Horizontal* confinement is essential to minimize energy losses from the resonance cavity by radiation, allowing high-Q designs. Horizontal confinement is achieved through reflections at the ends of the main phononic waveguide cavity. Such reflections arise due to a systematic mismatch between the acoustic impedance and dispersion characteristics of the main cavity waveguide and its boundaries. A suitable termination is required to provide such mismatch in a way that maximizes the specular reflections with negligible scattering.

As discussed in §3.1.2, the dispersion relation of the aforementioned phononic waveguides can be engineered by controlling their periodicity among different dimensions. The cavity termination can thus be implemented as a waveguide similar to the main cavity waveguide, while incorporating different periodicity or dimensions to produce a mismatch in the dispersion characteristics. Figure 4-5 shows a comparison between the dispersion relation of a main cavity waveguide and a possible termination waveguide. An arbitrary 25% larger period has been selected for the termination waveguide, leaving the main cavity mode in a partial bandgap of the termination waveguide. It is important to note that the larger period for the termination was preferred to shift the dispersion relation down, causing the main cavity mode to be above the *entire* termination mode curve and hence avoiding coupling to it. A termination with smaller period than the main cavity might be possible, but there is a big probability of coupling to the same mode at different β in the termination. This is also justified, as the main cavity and termination mode will have similar fields distribution.

When compared to abrupt termination of the main cavity, a termination based on a similar waveguide with mismatched dispersion characteristics represents smaller perturbations to the cavity structure. This provides the desired reflections without inducing significant scattering. It is also important to note that the phononic waveg-

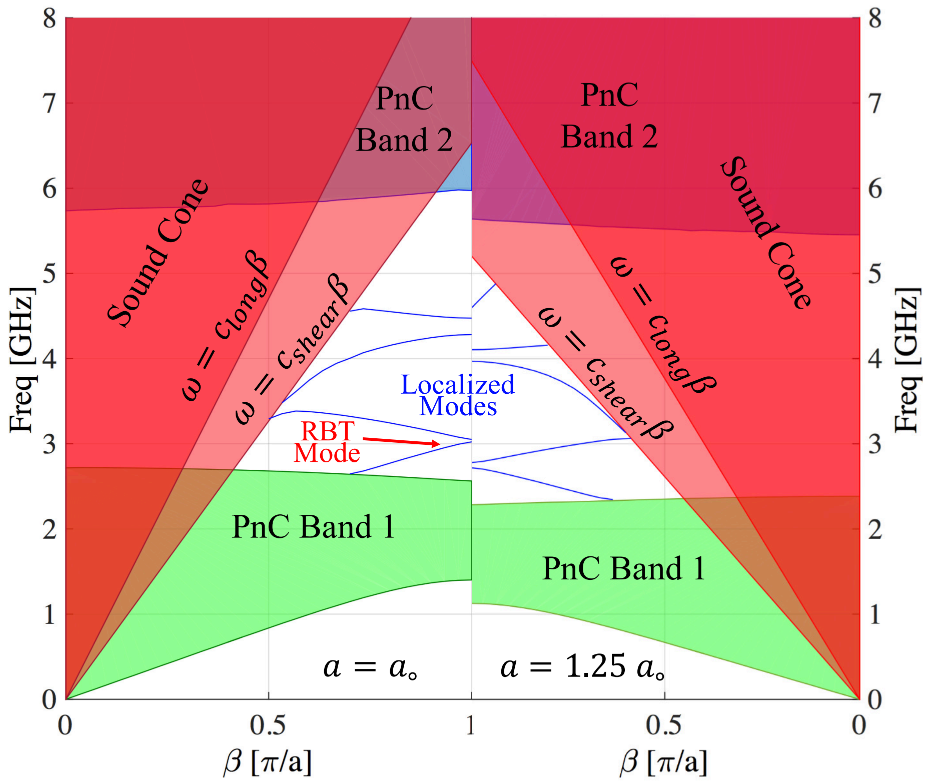


Figure 4-5: Comparison between the dispersion relation of a main cavity waveguide (right) and a possible termination waveguide with 25% larger period (left). The main cavity waveguide mode falls into the partial bandgap of the termination waveguide.

uides have discrete modes, and when operating far from the sound line, scattering amplitudes are significantly decreased in general (since there is no close by modes in the $\vec{k}$-space at the same frequency that can easily couple to the main cavity mode). Moreover, employing a waveguide for termination has the advantage that it provides the ability to precisely engineer the termination in order to optimize the quality factor of the RBT cavity. Also, for minimal scattering, an *adiabatic* transition (as discussed in §3.5) can be used, allowing a gradual and slow change of the waveguide period between the main cavity and the termination. In general, numerical optimization is required to find the best termination dimensions for minimal scattering.

4.1.4 Electrical Isolation

Another important factor in the RBT design is the reduction of the feed-through from input to output. There are multiple paths for feed-through in RBTs, as depicted in the equivalent circuit of Figure 4-6. While in passive MEMS resonators direct feed-through and leakage (represented by C_f and r_f) are the dominant factors, in CMOS-RBTs feed-through to the gate and body (through C_{fg} and C_{fb}) can be more detrimental. The problem with this type of feed-through is that it goes through the transistor electrical g_m and g_{mb}, which can be relatively high when compared to the electromechanical transconductance g_{em}. Thus, it is important to provide reliable electrical isolation between the drive MOSCAPs and sense transistors in the RBT to minimize feed-through.

The first thing to note in the equivalent circuit of Figure 4-6 is that a small resistance to RF ground (the DC biasing sources) from the transistor gate and body will reduce feed-through. To further mitigate the feed-through to the sensing MOSFET gate, a grounded gate is included between the drive and sense gates to screen the fields from drive to sense. The grounding vias of the MOSCAPs also help in shielding the sense transistor from the strong driving fields. Rectangular wall-like vias can be more effective in that respect.

It is equally important to reduce the feed-through to the transistor body. In SOI technologies like the IBM 32 nm SOI under consideration here, the driving MOSCAPs

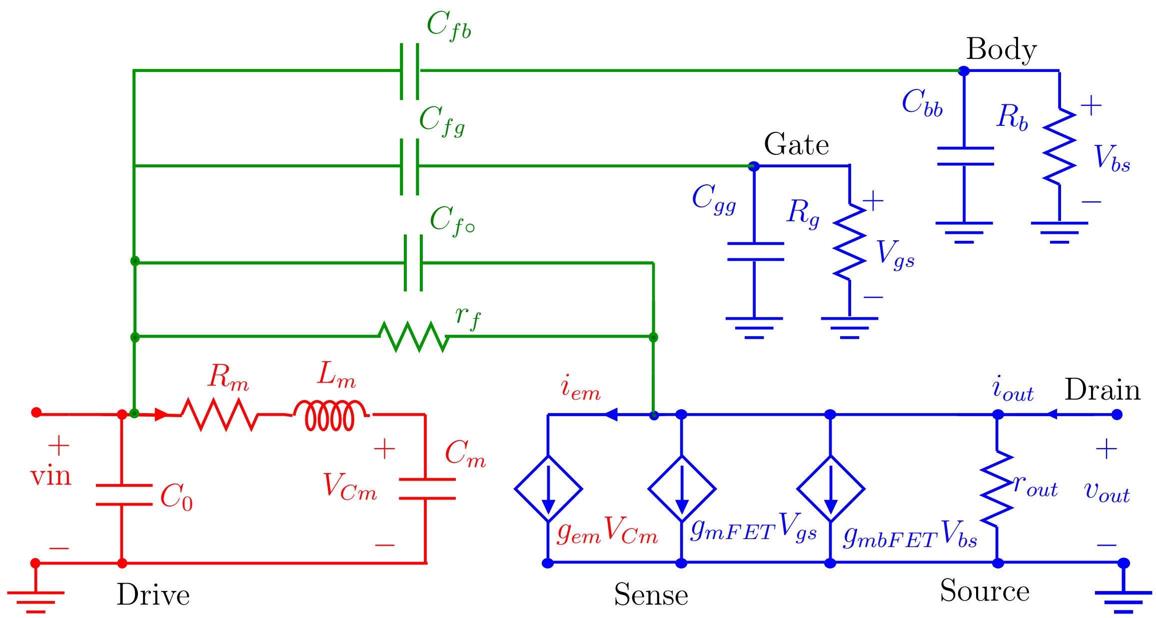


Figure 4-6: Equivalent circuit of the CMOS RBT, showing the different feed-through paths to the gate and body of the sense MOSFET.

and sense transistors can have separate active areas. This further helps in reducing feed-through to and from the bulk wafer. The thin SOI layers allow such isolation without major disruption to the cavity. RBT design optimization will fine-tune the isolation and sensing gate lengths to minimize the scattering at this isolation. Such an isolation technique might not be possible in bulk CMOS technologies; as incorporating a shallow trench isolation (or STI) between the drive MOSCAPs and sense transistors usually results in major cavity disruption because STI is relatively thick compared to the FEOL layers that form the waveguide. Also, STI usually comes with large DRC restrictions, making it difficult to optimize the cavity for its inclusion.

In general, isolation gates and low-resistance contacts with bypass capacitors to the gate and body of the sensing MOSFET should effectively reduce the feed-through. Also, fully differential driving helps reduce electrical feed-through to the bulk as the fields tend to cancel each other.

4.1.5 Full RBT Structure

The full RBT structure incorporates all the previously discussed concepts. Figure 4-7 shows the left-half of the full RBT structure, which is anti-symmetric about

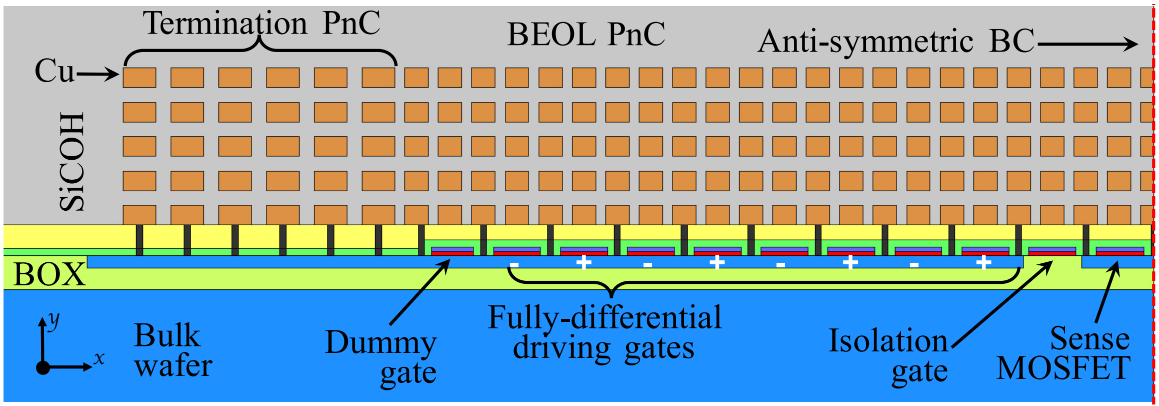


Figure 4-7: Left half of an actual RBT cavity structure with vias in the termination waveguide. The structure is anti-symmetric around the right axis.

the y-axis. Anti-symmetry represents a symmetric geometry with anti-symmetric loads. The structure clearly shows the main RBT cavity with fully-differential driving MOSCAPs. The termination waveguide section is also included to provide horizontal confinement. Sensing MOSFET with separate active area from the driving MOSCAPs is shown, with a grounded isolation gate in between. Although the thin FEOL layers result in minor disturbance to the cavity structure, full numerical optimization is required to match the waveguide characteristics and reduce scattering as this discontinuity.

4.2 RBT Structural Optimization

As discussed in §4.1, numerical optimization is necessary to match the different waveguide characteristics to achieve the least scattering and hence the best quality factor. Generally, maximizing RBT performance through numerical optimization is highly desirable; however RBT simulations are numerically demanding, which prohibits numerical optimization with reasonable run-times. This section demonstrates the implementation of different numerical techniques with the goal of speeding up calculations and making numerical optimization feasible.

4.2.1 A Numerically Intensive Problem

Numerical optimization is usually concerned about minimizing (or maximizing) a certain cost function or objective function. The objective function for RBTs can be the resonator quality factor, output signal amplitude, or any resonator figure of merit. Most optimization techniques require the evaluation of the objective function (often with its gradient) many times for different parameter sets. Evaluation of the objective function has to be numerically efficient for the optimization process to be completed in a timely manner.

Radiation Modes

Since RBTs incorporate a plethora of materials and hundreds of interfaces, analytical estimation of their performance is prohibitively difficult, calling for numerical simulations. FEM simulation is generally the option of choice for simulating such complicated structures, which naturally results in a numerically demanding problem. It is also important to notice that unlike *released* MEMS resonators with free boundaries, unreleased RBTs are completely encapsulated in solid materials. Radiation into the material continuum surrounding the resonator is often the major limitation of the resonator quality factor Q. This complicates FEM simulations as the domain surrounding the resonator has to be modeled to simulate such losses. Perfectly matched layers (PMLs) are used in this case to emulate radiation losses by allowing plane waves propagation *without reflections* outside the resonator domain, at all frequencies [77, 83].

For released resonant MEMS structures with free boundaries, eigenmode analysis is the most convenient way to characterize the resonator. It is numerically efficient to find the eigenvalues (i.e., eigenfrequencies) of a matrix that are close to a certain value. Resonator performance estimation in this case can be performed quickly and efficiently. However, the unreleased nature of the RBTs considered here, with the inclusion of PMLs all around the resonator, renders eigenmode analyses ineffective. This is justified by considering the dispersion relation of Figure 4-2. Since the RBT structure does not incorporate translational symmetry, $\vec{k}$ is not conserved in the structure. As a result, eigenmode analysis will detect, at each frequency, all possible modes for all $\vec{k}$. Each frequency corresponds to a horizontal line in Figure 4-2, with modes that can be identified as

- all plane waves in the bulk wafer, corresponding to the longitudinal and shear sound cones;
- all plane waves in the dielectric above the structure, corresponding to much shallower sound cones than those of silicon; and
- the actual resonator confined modes, and
- surface modes outside the resonator cavity.

The direct implication is that a large number of eigenmodes is required to isolate the confined resonance mode from all the radiation modes, making the problem numerically intensive.

Frequency Domain FEM Simulation

In this case, small signal FEM frequency domain analysis is a better alternative to eigenfrequency analysis. Frequency domain analysis involves sweeping the frequency of a small signal input source to find the resonator transfer function from input to output. For simplicity, the resonator input is considered to be stress loads with the correct phase applied to the resonator driving MOSCAPs, whereas the average stress in the sensing transistor channel is used as output. Frequency domain analysis captures only the modes which can be efficiently driven and sensed, *and* induce mechanical resonance.

In some applications, maximizing the resonator quality factor Q becomes the main goal with relaxed constraints on the exact frequency specifications. Clocks for digital circuits or a reference source for a phase-locked loop (PLL) are good examples. Such constraint relaxation becomes a numerical challenge for optimization based on frequency domain simulation. As the resonator dimensions get varied in the course of optimization, the resonance frequency can change significantly. A wide frequency sweep (over a bandwidth $B_\circ$) is thus required in the frequency domain analysis to follow the resulting resonance frequency variation. Moreover, with the goal being to maximize Q, the resonance peak becomes narrower in the frequency domain. Wide frequency range and high resolution requirements result in a dramatic increase in the number of simulation frequency points required to resolve the resonance peak. Simulation times become prohibitively long for practical optimization purposes.

This section provides an efficient solution for this problem. The presented technique is based on model order reduction (MOR) to significantly speed-up frequency domain simulations. Memoization is also used to store previous simulation results, allowing a prediction of the resonance frequency as the optimization proceed, and hence limiting the simulation band required. Finally, gradients are evaluated separately by finite differences over very narrow frequency ranges, allowing for even greater increases in speed.

4.2.2 Problem Formulation

The RBT optimization problem can be formulated as

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & f(\boldsymbol{x}) \\ \text{s.t.} \quad & \boldsymbol{c}(\boldsymbol{x}) \leq 0 \\ & 0 \leq x_i \end{aligned} \tag{4.9}$$

where $\boldsymbol{x}$ is the resonator N-dimensional geometrical parameters vector, $f(\boldsymbol{x})$ is the objective function, and $\boldsymbol{c}(\boldsymbol{x})$ is a non-linear constraint function. The resonator parameters are normalized to their initial guess; hence all parameters are on the order of 1. This normalization greatly simplifies tolerances and step size calculation.

Objective Function

The objective function $f(\boldsymbol{x})$ is a scalar measure that can be chosen to maximize the resonator Q, to maximize electromechanical transconductance g_{em}, to minimize spurious modes, or any other resonator figure of merit. For example, to maximize the quality factor, one would select $f(\boldsymbol{x}) = -Q(\boldsymbol{x})$. This objective function will favor designs with the best energy confinement. Another objective function can be $f(\boldsymbol{x}) = -|g_{\mathrm{em}}(\boldsymbol{x})|$ to maximize the device g_{em}. In general, the two objective functions may result in different designs. This can be explained by considering the RBT as a cascade connection of three independent systems: driving transducers, mechanically resonating structure, and sensing transducer. Maximizing Qonly requires minimizing the losses in the resonant cavity. On the other hand, maximizing the transconductance g_{em}requires maximizing the stress at the transducers, which can be achieved by specific mode shapes*.

It is important to note that when formulating the problem, there is no need to go into the specific details about matching the characteristics of the different waveguide sections in the resonator. The optimizer will select the design that optimizes the objective function. The requirements to match the waveguide characteristics are implicit in demanding the best resonator performance and will be naturally satisfied in the course of optimization.

DRC Constraints

A challenging aspect of RBT design is to remain compliant with the CMOS design rule check (DRC) constraints. DRC constraints usually involve rules about the allowable gate lengths and separation, metal widths and separations, metal filling densities, and so on. These constraints are imposed by the CMOS foundry to guarantee successful fabrication of the devices with sufficient yield, or sometimes to protect the fabrication facility itself†.

In the optimization problem (4.9), the non-linear function $\boldsymbol{c}(\boldsymbol{x})$ corresponds to the DRC constraints. The reason for non-linear constraint representation is that DRC constraints are usually discontinuous. For example, when metal lines reach a certain

*In traditional MEMS resonators, such distinction is not very clear as the mode shapes are usually optimal in terms of transduction.

†Some DRC rules are imposed to prevent different materials delamination and lift-off, which might contaminate subsequent processing machinery.

threshold width, the DRC-required separations may increase.

Assume the designed separation between metal lines on layer i to be s_i and the DRC required value to be $s_i^{\mathrm{DRC}}(w_i)$, where w_i is the width of the metal line. The corresponding constraints element $c_i(\boldsymbol{x})$ is given by

$$-\left(s_i - s_i^{\mathrm{DRC}}(w_i)\right) \leq 0. \tag{4.10}$$

This formulation is adopted so that the actual value of $c_i(\boldsymbol{x})$ is proportional to the DRC constraints violations. This will help most optimization algorithms to efficiently choose the next design point, as $\boldsymbol{c}(\boldsymbol{x})$ appears directly in the Lagrangian and KKT conditions as $\Lambda^T \cdot \boldsymbol{c}(\boldsymbol{x})$, where Λ is a vector of Lagrange multipliers λ_i [101, 102].

It is important to note that the resulting Lagrange multipliers λ_i from the optimization are very useful from a CMOS process design point of view. If certain DRC constraints are *tightly* satisfied ($c_i(\boldsymbol{x}) = 0$), the corresponding Lagrange multipliers will be non-zero. These non-zero multipliers also indicate which constraint has the greatest effect on the objective function. Such detailed information can be used by process engineers who may consider relaxing specific DRC rules or optimizing the process differently for better resonator performance.

4.2.3 Objective Function Evaluation Speed-Up

The Problem of Long Simulation Times

During the course of optimization, the objective function needs to be evaluated several times, sometimes hundreds or even a thousand times (depending on the number of variables). With the problem at hand, each objective function evaluation involves a full frequency domain FEM simulation for the entire RBT. Such simulation is numerically intensive and becomes more demanding for complicated geometries such as the unreleased RBTs. Typical RBT designs considered here averaged 40 minutes for a single evaluation. In order to speed up the objective function evaluation, it is important to understand where most of the simulation time is spent.

First of all, it should be noted that the FEM simulation offers a complete solution for the entire domain. The resulting solution includes the displacement at each and every point in space (even in the PML), for all frequencies. However, the optimization is problem only concerned with the resonator transfer function from input and output. This means that only the average stress in the sensing transistor channel is of interest.

Next, as mentioned above, frequency sweep over a wide frequency range with many points is needed in order to capture the high-Q peak of the resonator. The resonator response, however, is negligible for the most part of this wide frequency band, except for any existing spurious modes. Most of the time is spent solving the full FEM problem, many times, for frequencies and results that the optimizer does not care about. Furthermore, away from the main peak and spurious modes, the resonator response does not change rapidly and can be easily predicted.

MOR by AWE and Padé Approximants

From the above discussion, the problem at hand is a good candidate for model order reduction (MOR). A simple rational transfer function approximation with few pairs of complex-conjugate poles should be sufficient to evaluate $f(\boldsymbol{x})$, without the need for the full FEM solution for every point in space at each frequency. Multiple model order reduction techniques can be used to extract the rational transfer function approximation. However, COMSOL readily includes model order reduction that can be incorporated by enabling the asymptotic waveform evaluation (AWE) feature in the COMSOL solver [76].

COMSOL AWE implementation calculates low-order Padé approximation [103] for a given output (in this case the average sensing channel stress) on small frequency intervals. The Padé approximation of type p/q for the frequency response $H(\omega_\circ + \sigma)$ around frequency $\omega_\circ$ is given by

$$H_{p,q}(\omega_\circ + \sigma) = \frac{b_p\sigma^p + \cdots + b_1\sigma + b_\circ}{a_q\sigma^q + \cdots + a_a\sigma + a_\circ}. \tag{4.11}$$

Its Taylor series about $\sigma = 0$ matches that of $H(\omega_\circ + \sigma)$, at least for the first $p + q + 1$

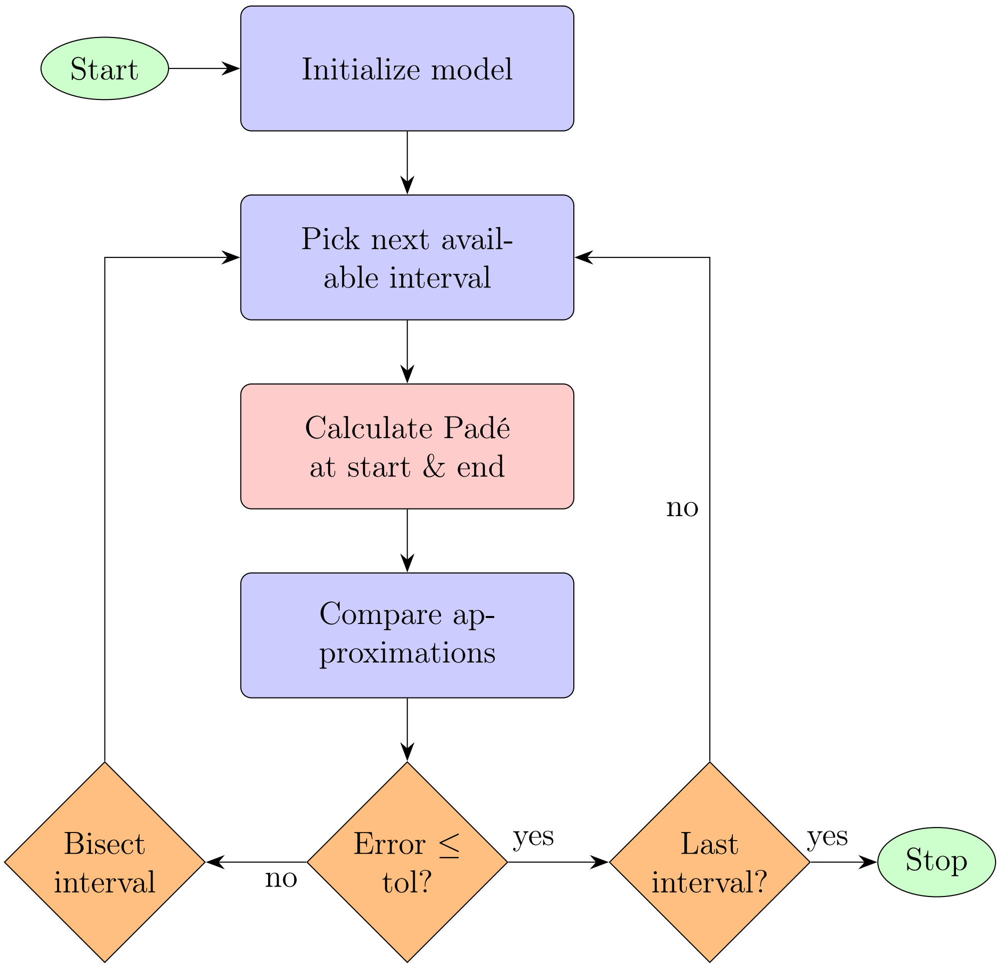


Figure 4-8: Flow chart for Padé approximation calculation in COMSOL.

terms, as

$$H_{p,q}(\omega_\circ + \sigma) = H(\omega_\circ + \sigma) + \mathbb{O}(\sigma^{p+q+1}). \tag{4.12}$$

The q^{th} order Padé approximant $H_q(\omega_\circ + \sigma)$ is found by assuming that $p = q - 1$. In AWE, $H_q(\omega_\circ + \sigma)$ is found by calculating the leading $2q$ moments of $H(\omega_\circ + \sigma)$ as described in [103].

A flowchart of the AWE simulation procedure in COMSOL is shown in Figure 4-8. It starts with a given frequency interval, where a Padé approximation is calculated both at the start and end of the interval. Next, the result of both approximations is evaluated and compared at several points in that frequency interval. If the results match within a certain tolerance, the interval is accepted; otherwise, the interval is bisected and the process repeats.

This technique is very efficient with the wideband sweeps under consideration. Large intervals can be used in the smooth regions away from the peaks allowing for very fast simulation in these regions. When a resonance is encountered, the simulation will automatically drop to smaller intervals until the simulation tolerance is met. The higher the order of the Padé approximation, the larger the intervals that will be accepted. However, higher order Padé approximations are not generally useful beyond $q > 8$. This is mainly due to the ill-conditioning of the moments matrices [103]. Padé approximations with $q = 5$ were found to provide the most useful speed-up. A typical speed-up of $8\times$ was observed when employing AWE in these simulations.

In case of multiple resonance peaks appearing in the objective function, only the one with the best objective value is selected. If the optimization is targeting eliminating the spurious modes within a certain bandwidth, such designs can be subsequently penalized in the objective evaluation. All optimizations considered here were focused mostly on maximizing the quality factor and transduction, while allowing for spurious modes.

Memoization

During optimization, the optimizer may suggest points that are close in the design parameter space. For such points, the resonance frequencies are not expected to be very different from each other and there is no need to simulate the full bandwidth $B_\circ$. This *a priori* knowledge can be exploited to further provide more speed-ups to the objective function evaluation.

For this reason *memoization* is incorporated into the objective function evaluation. The resonance frequency, quality factor, and different performance metrics for each design point simulated are saved as a *side effect* of the objective function evaluation*. Whenever the objective function value is needed for a new design point, all previous points are searched for the nearest point in the design parameter space. If the latter falls within a maximum Euclidean distance, this point is used to estimate the new design resonance frequency. With the optimization design parameters representing

*The objective function $f(\boldsymbol{x})$ remains idempotent under sequential composition $f;\, f$.

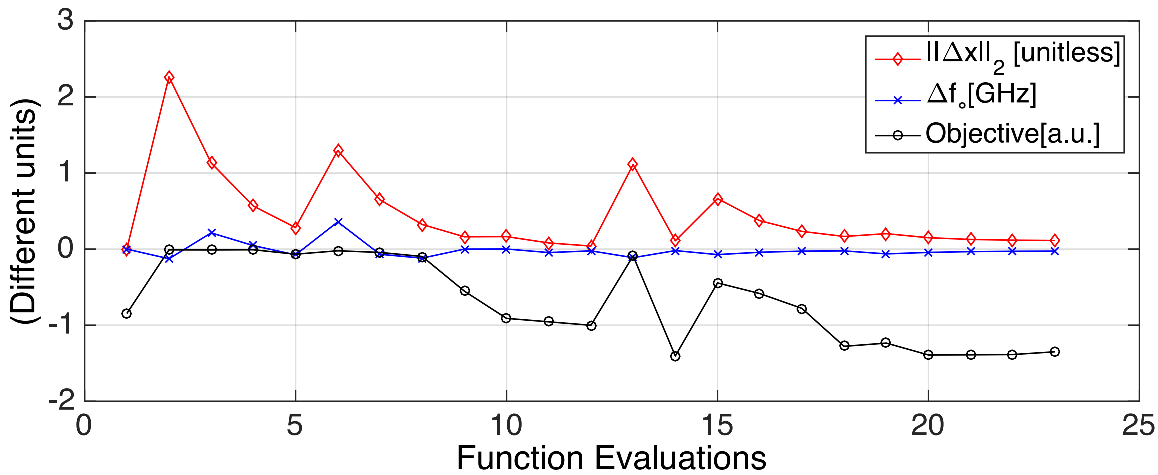


Figure 4-9: Euclidean distance ($||\Delta x||_2$) and resonance frequency change ($\Delta f_\circ$) from starting design for different function evaluations.

physical device dimensions*, the resonance frequency is expected to change at most linearly (likely sublinear) with the parameters as in §2.1.6. In this case, a limited frequency sweep over a bandwidth $B_i << B_\circ$ is used for much faster FEM simulation.

It is also interesting to note that as the optimization converges, the selected points become closer in the design space. This is shown in the example run data of Figure 4-9. In this case, the optimization benefits greatly from memoization and the objective function evaluation gets faster as the optimization progresses.

Major speed-ups are achieved through memoization on the order of $B_\circ/B_i$. In fact, only a few points of the design space end up being evaluated over the full frequency band $B_\circ$.

Gradient evaluation

Another important aspect for optimization speed-up is the quick evaluation of the gradients of both the objective function and constraints. The constraints in (4.10) are just analytical functions of the design dimensions, hence their gradient evaluation does not limit the optimization runtime. This is not the case for the objective function that is evaluated by FEM simulations. The objective function is non-linear with no

*Parameters representing the number of termination waveguide periods, for example, are not considered physical dimensions.

analytical form available. Finite difference becomes the option of choice in this case, with the i^{th} component of the gradient evaluated as

$$(\nabla f)_i = \frac{f(\boldsymbol{x} + \epsilon\,\hat{\boldsymbol{e}}_i) - f(\boldsymbol{x})}{\epsilon}, \quad \forall i \in [1, N], \tag{4.13}$$

where $\hat{\boldsymbol{e}}_i$ is a unit vector with only the i^{th} component set to 1.

Objective function gradient evaluation requires N FEM simulation for each design point. However, since the wave equation is linear under coordinate scaling (§2.1.6), the resonance frequency $f_\circ$ can at most change by* $\pm\epsilon$. Thus, it is sufficient to consider a frequency bandwidth $dB_i = 2\,\epsilon\, f_\circ^i + 2\,BW^i$, with $BW^i = f_\circ^i / Q^i$ for gradient estimation. Choosing $\epsilon = 0.5\%$, the gradient simulation bandwidth dB_i is usually much smaller than the full problem frequency band $B_\circ$. The speed-up in gradient evaluation S_∇^i compared to a naïve finite differencing is given by

$$S_\nabla^i = \frac{B_\circ}{2\,\epsilon\, f_\circ^i + 2\,BW^i} = \frac{B_\circ}{2\epsilon\, f_\circ^i} \times \frac{Q^i}{Q^i + 1/\epsilon} \approx \frac{B_\circ}{2\epsilon\, f_\circ^i}, \tag{4.14}$$

which is typically large, even if $B_\circ$ is only a few percents of $f_\circ$.

4.2.4 Optimization Flow

The complete optimization flow incorporating all the previously discussed concepts is shown in Figure 4-10. The MATLAB constrained optimization function `fmincon` with an interior-point algorithm was used for this problem [101]. Objective function evaluation relied on frequency domain FEM simulation with COMSOL Multiphysics [76].

The optimizer in MATLAB generates a design point $\boldsymbol{x}^i$ at each iteration i. To evaluate the objective function, a search is first performed over all previous design points to find the nearest one, as intended by the memoization technique. A simulation bandwidth B^i is selected based on resonance frequency estimation from memoiza-

*The actual sign of the frequency change depends on the dimension. If the dimension tends to make the cavity smaller, say the length of termination within a limited area, the resonance frequency may increase with ϵ. However, for actual cavity dimensions, the frequency is likely to decrease with ϵ.

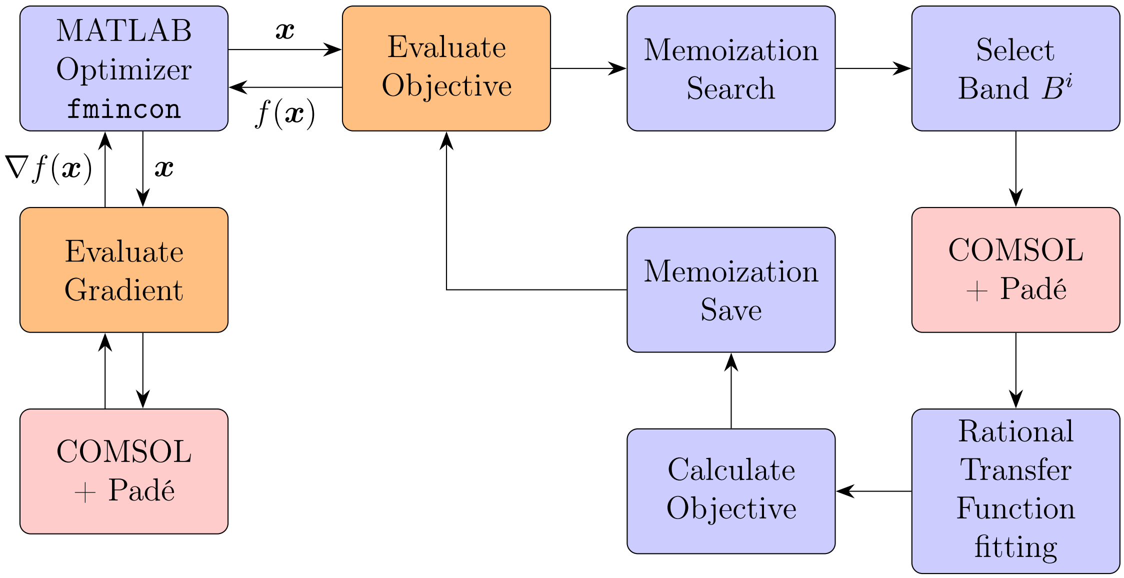


Figure 4-10: Implementation of the proposed optimization flow.

tion. COMSOL Multiphysics frequency domain simulation is called from MATLAB. AWE model order reduction by Padé approximations is enabled in COMSOL itself. The simulation returns the resonator frequency response. A rational transfer function fitting using MATLAB's `ratoinalfit` function is performed to accurately find $f_\circ$, Q and the peak transconductance. The objective function is calculated and the memoization state is saved for future evaluations. The gradient evaluation, when requested, proceeds simply as described above.

4.2.5 Optimization Flow Application

All RBTs considered in this chapter have been optimized with the previous flow, and the resulting designs are reported in subsequent sections. The structure of Figure 4-7 is considered as an example to illustrate the optimization flow performance with 6 design parameters. Starting from a close enough initial guess, the optimization completes (achieving double the starting Q and 50% larger $|g_{\mathrm{em}}|$) with 21 function and gradient evaluations in 4 hours, marking a 5× improvement over naïve gradient evaluation and 40× improvement when not using the proposed framework*. The small speed-up in gradient evaluation is mainly due to rebuilding the geometry in

*The 40× improvement factor is estimated based on improvement in single evaluation.

COMSOL before every evaluation. The resonator geometry is fairly complex and involves multiple domains to the point where the time required by COMSOL to construct the geometry is significant*. In general, an arbitrary starting point will require more iterations and longer time; hence the benefit of the presented technique.

The presented technique can be easily generalized to the optimization of any MEMS device that requires frequency domain simulation. It is also important to note that the proposed optimization flow can be used for CMOS process characterization purposes. While actual layer thicknesses can be accurately determined by focused-ion-beam (FIB) etching and SEM/TEM imaging, the mechanical properties of the layers *as in the final die* are hard to measure. By using the exact dimensions from FIB images, with material properties as model parameters, the above-mentioned optimization flow can be used to efficiently extract these properties by matching FEM simulation results to measured resonator performance.

4.3 RBT Implementations

Multiple RBT structures have been implemented based on the concepts of §4.1. This section demonstrates different RBT implementations, with simulated and measured results when available.

4.3.1 RBT with PnC Waveguide Termination

Multiple RBTs have been implemented and optimized for IBM 32 nm SOI technology. A cross section of the first structure, which we will refer to as "RBT-A," is shown in Figure 4-11. Another similar structure, referred to as "RBT-B," is shown in Figure 4-14.

In both RBTs, the main RBT cavity is formed from a phononic waveguide like that described in §3.1. Another waveguide with a larger period is used for termination, with a single dummy gate included at its start. Four fully differential MOSCAPs

*A single dimensional parameter update may require building the entire structure, as all the waveguide sections are affected.

Table 4.1: Optimization dimensional parameters for "RBT-A" and "RBT-B". Dimensions are in nanometers.

Parameter	RBT-A	RBT-B	Description
l_{M}	154	154	Length of the main cavity PnC metal stripes
s_{M}	66	66	Separation of the main cavity PnC metal stripes
$l_{\mathrm{M\text{-}T}}$	214	214	Length of the termination PnC metal stripes
$s_{\mathrm{M\text{-}T}}$	98	98	Separation of the termination PnC metal stripes
l_{g}	314	314	Length of the sensing transistor gate
l_{iso}	312	312	Length of the isolation gate

are used for driving on each side of fully differential sensing transistors. Isolation gates are also included to help reduce feed-through. The difference between "RBT-A" and "RBT-B" is that the latter includes contacts to active area (`CA` contacts in IBM 32 nm) in the termination waveguide section.

Full numerical optimization has been performed for both RBT structures. Maximizing the transconductance has been selected as the objective function for the optimization. Artificial isotropic material losses that exceed the physical intrinsic losses have been included in the FEM models. This sets a limit of 3,000 on the quality factor in simulation. Such a limit is intended to avoid generating spurious sharp resonances that can disrupt the optimization process. The optimization usually concludes with RBTs having Q of 3,000, the artificial material losses limit. This indicates that radiation losses have been reduced to guarantee a quality factor of at least 3,000. A lower limit on the quality factor is usually a good specification for resonators for low-phase noise timing applications. The fabricated resonators are estimated to have higher quality factors.

The optimization parameters, along with the final design values, are listed in Table* 4.1. FEM simulation results are reported for the optimized structures. The simulation assumes a 1 MPa T_{yy} stress applied at the driving MOSCAPs with the proper sign to emulate fully differential driving. Quality factor was artificially limited in this simulation to $3,000$. The average stresses at the sensing transistors for "RBT-B" are shown in Figure 4-15, while the T_{yy} stress at resonance is shown in Figure 4-16.

*The devices have actually been optimized with a different crystal orientation for the bulk wafer, resulting in 15% change in its stiff matrix values. The presented designs are in fact suboptimal.

It is clear that the structure achieves almost perfect confinement, both in the vertical and horizontal directions, without significant scattering anywhere in the structure. This is a direct result of the full structural optimization that implicitly matched the characteristics of the different waveguides to achieve high Q.

4.3.2 RBTs with Adiabatic Transition

As discussed in §3.5, adiabatic transition can be used for RBTs terminations. Adiabatic transition is intended to reduce scattering in the RBT termination.

A spatial taper function $\phi(\zeta)$ that is infinitely differentiable is given by

$$\phi(\zeta) = e \exp\left(\frac{-1}{(2-\zeta)\zeta}\right) \quad , \quad \forall\, \zeta \in [0,1]. \tag{4.15}$$

For a taper that starts at $x = 0$ with a value a and ends at $x = 1$ with a value $b > a$, the taper function $f_{\text{Taper}}(\zeta)$ takes the form

$$f_{\text{Tapper}}(\zeta) = b + (a-b)\phi(1-\zeta) \quad , \quad \forall\, \zeta \in [0,1]. \tag{4.16}$$

This taper function is shown in Figure 4-17. Since the waveguide is periodic with discrete translation symmetry, waveguide periods are set to sample the taper function $f_{\text{Taper}}(\zeta)$ with equal ζ intervals.

Five termination waveguide periods have been selected to implement adiabatic transition. These were found to lower scattering while maintaining reasonable resonator footprint. The RBT structure, referred to as "RBT-C," is shown in Figure 4-18. Numerical optimization has been carried out for the RBT with adiabatic transition. Frequency response of the RBT as well as the T_{yy}-stress at resonance are shown in Figure 4-19 and Figure 4-20, respectively. It is clear that the mode shows minimal scattering.

The RBT with adiabatic transition shows a quality factor of 2,450, whereas an RBT with the same dimensions but with abrupt termination shows a quality factor of 1,820. That is, FEM simulation predicts that the adiabatic transition for the

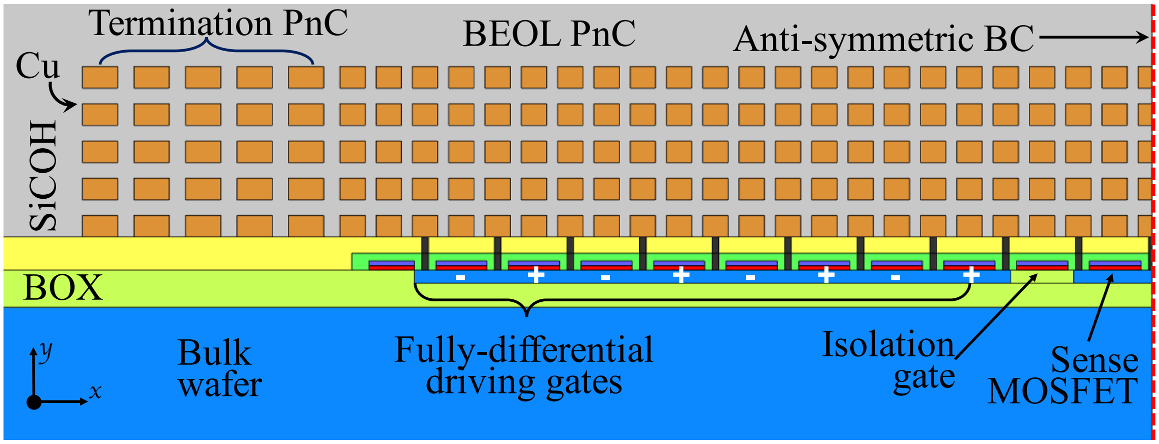


Figure 4-11: “RBT-A,” an RBT with PnC waveguide termination.

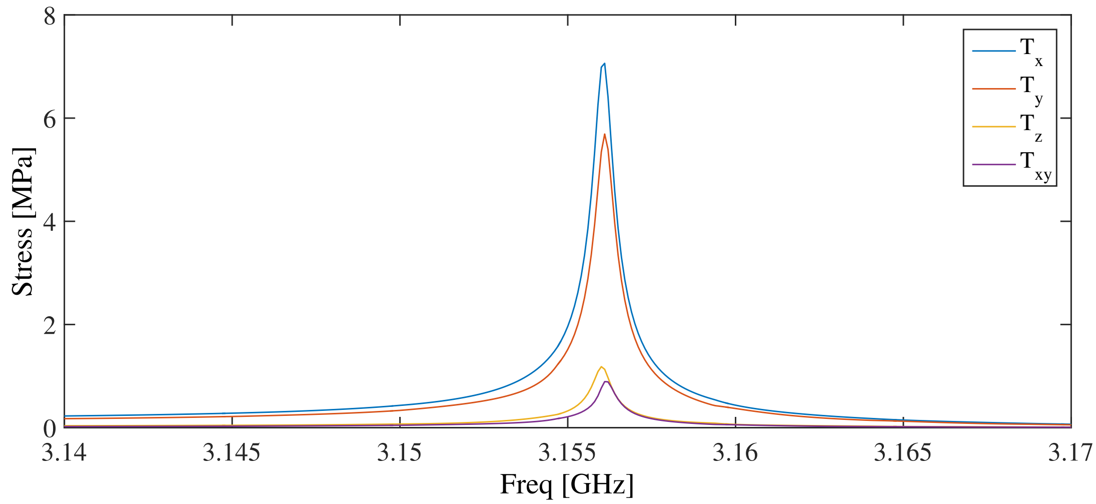


Figure 4-12: FEM simulation results showing the average stresses at the sensing transistors for “RBT-A” with a resonance at 3.155 GHz.

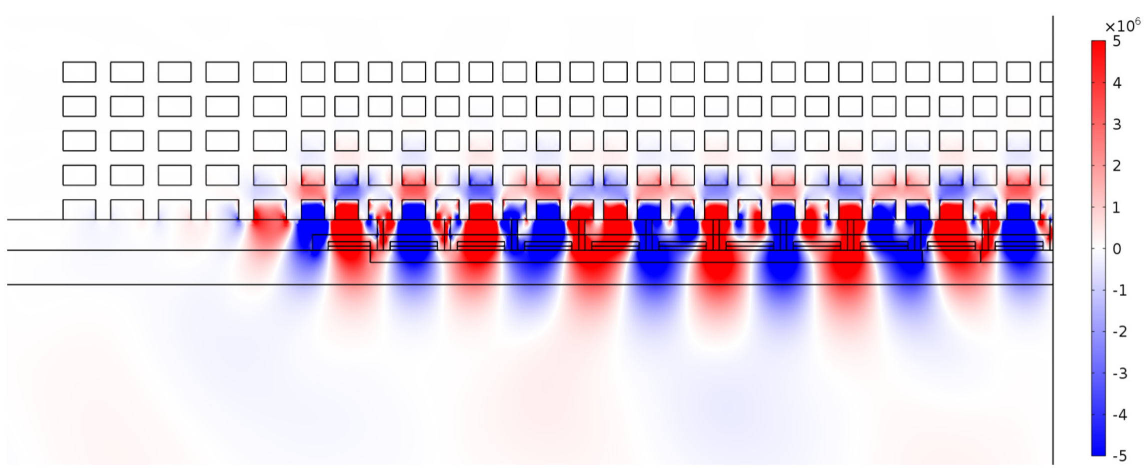


Figure 4-13: FEM simulation showing the y-stress T_{yy} for “RBT-A” at the 3.155 GHz resonance mode.

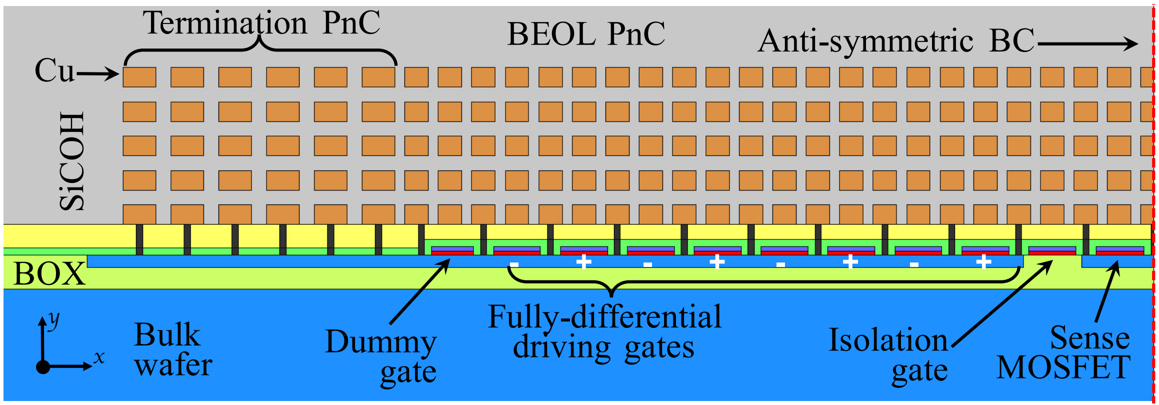


Figure 4-14: "RBT-B," an RBT with contacts in the termination waveguide.

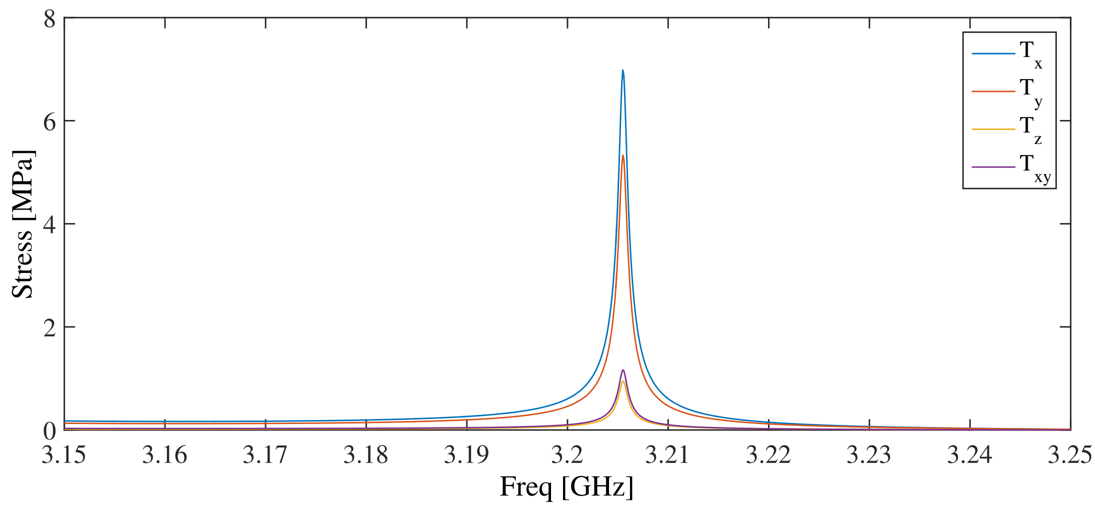


Figure 4-15: FEM simulation results showing the average stresses at the sensing transistors for "RBT-B" with a resonance at 3.205 GHz.

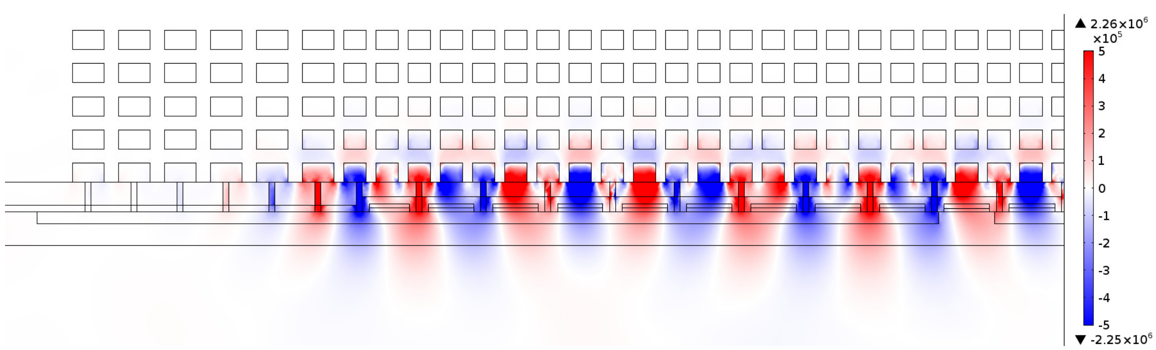


Figure 4-16: FEM simulation showing the y-stress T_{yy} for "RBT-B" at the 3.205 GHz resonance mode.

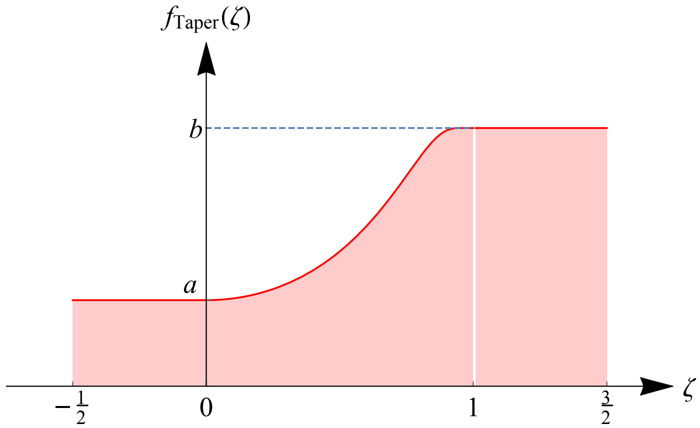


Figure 4-17: Taper function used for the adiabatic termination of the RBTs.

termination lowers the scattering and enhances Q by 35%. It is also important to note that the mode of Figure 4-20 is uniformly distributed among the driving and sensing. This becomes especially obvious when compared to the modes of Figure 4-13 and Figure 4-16. Thus, although the Q is not higher (due to the particular dimensions chosen for this design), the mode is quite uniform and results in efficient driving and sensing. This is directly evident by considering the stresses in Figure 4-19, which are much higher than those of "RBT-A" and "RBT-B," even though the overall Q is lower.

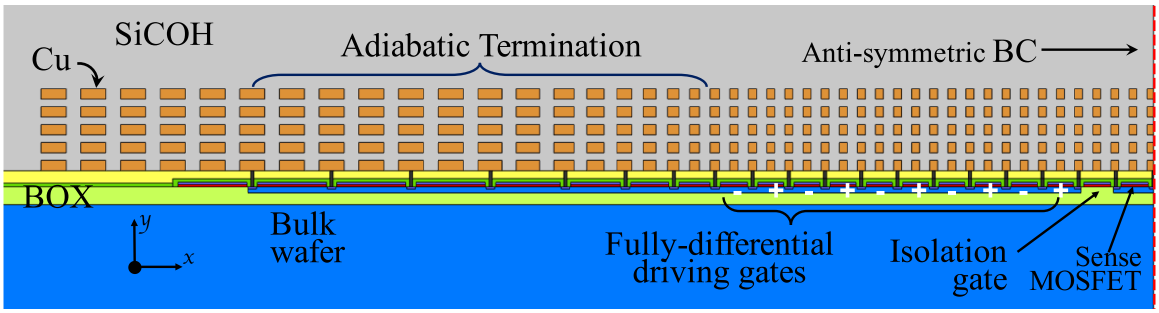


Figure 4-18: "RBT-C," an RBT with an adiabatic termination

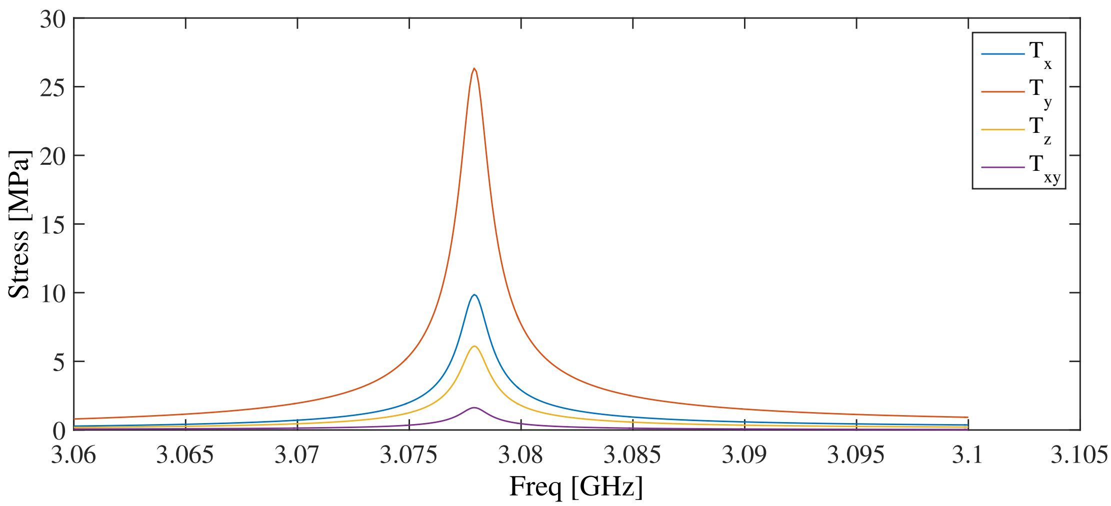


Figure 4-19: FEM simulation results showing the average stresses at the sensing transistors for "RBT-C" with a resonance at 3.078 GHz.

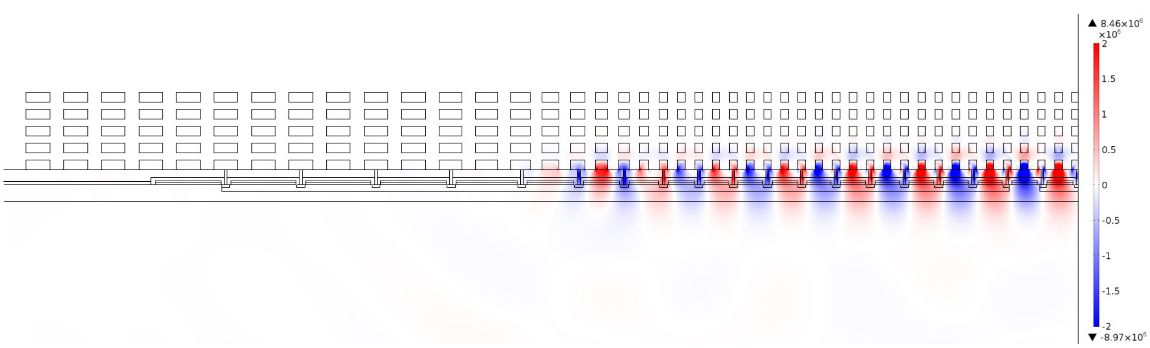


Figure 4-20: FEM simulation showing the y-stress T_{yy} for "RBT-C" at the 3.078 GHz resonance mode.

4.4 Fabrication in IBM 32 nm SOI

The RBTs of §4.3 have been implemented and fabricated in IBM 32 nm SOI technology, without any post-processing or any modification of the regular IBM process flow. This section describes the layout, test structures, and cross-sections of the actual fabricated devices.

4.4.1 General Layout and Implementation Considerations

SKILL ROD P-Cells

All RBTs under consideration have periodic regions of metal layers and gate stack. For accurate implementation, the Cadence SKILL language was used to create *parameterized cells* (P-Cells) for all the RBTs considered. P-Cell parameters include those of Table 4.1, among other dimensional parameters, for different routing and isolation. Termination, contacts, and adiabatic transition are also included as options in the P-Cell, along with wall-like rectangular vias and regular square vias. This approach resulted in a single SKILL P-Cell that automates the implementation of all RBTs considered.

The P-Cell makes extensive use of Cadence relative object design (ROD) in the SKILL language to significantly simplify the positioning of the different RBT elements. Moreover, the non-designed dimensions (like via size, transistor gate separation, active area extension, etc.) are all directly inferred from the technology file rules by the P-Cell. This guarantees that the P-Cell always follows the PDK design rules (given that the requested metal widths and separation comply with such rules). It also significantly simplifies the migration of the RBT P-Cell to other technologies when the need arises.

Testing Pads

All implemented RBTs rely on fully differential drive and sense transduction. A balanced-balanced testing structure is adopted. Two sets of GSSG RF pads with 100 μm pitch are used for drive and sense, respectively. The full pads structure is

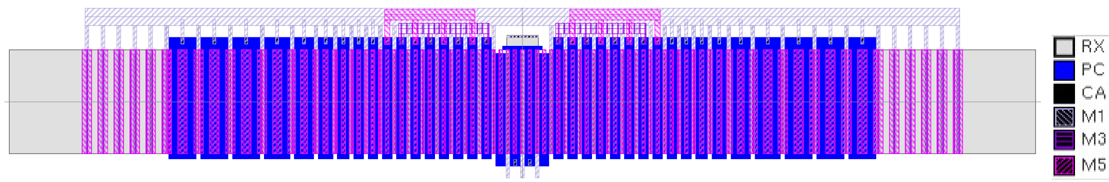


Figure 4-21: P-Cell layout used to generate the different RBTs considered in this work.

shown in Figure 4-22. It minimizes the routing to the resonator terminals, resulting in reduced parasitics, which later simplifies de-embedding. Moreover, grounded metal lines and vias are included between the drive and sense GSSGs on all available metal layers. This is intended to act as *shielding* for sensing and thus minimize the direct feed-through between drive and sense. All ground pads are connected to each other and to (i) the source of the sensing MOSFETs, (ii) the body of the sensing transistors, and (iii) the body and channel of the driving MOSCAPs.

A separate DC pad is included to provide gate biasing for the sensing transistor. Minimizing noise on this gate bias, feed-through, and resistance to RF ground is essential for maximizing the RBT output signal-to-noise ratio as discussed in §4.3. A large 11 pF vertical natural capacitor (VNC) is included under the DC pad itself. This large VNC serves as a bypass capacitor from gate biasing pad to ground. The gate bias signal is then routed to the actual RBT sensing gate through a metal line shielded from all directions to minimize feed-through from the RF signals involved in testing. Large MOSCAPs to ground are also connected just before reaching the actual sense transistor gate, in order to further reduce the noise on this DC bias signal.

ESD Protection

As the RBT under test is in direct contact with the testing pads, it is important to protect the former from ESD events. The terminals directly connected to the pads are the sensing transistor drains, the sensing transistor gate, and the MOSCAP driving terminals. The sensing transistors drains are connected to a sufficiently large grounded active area (the body of the sensing FET), and hence there is no need for

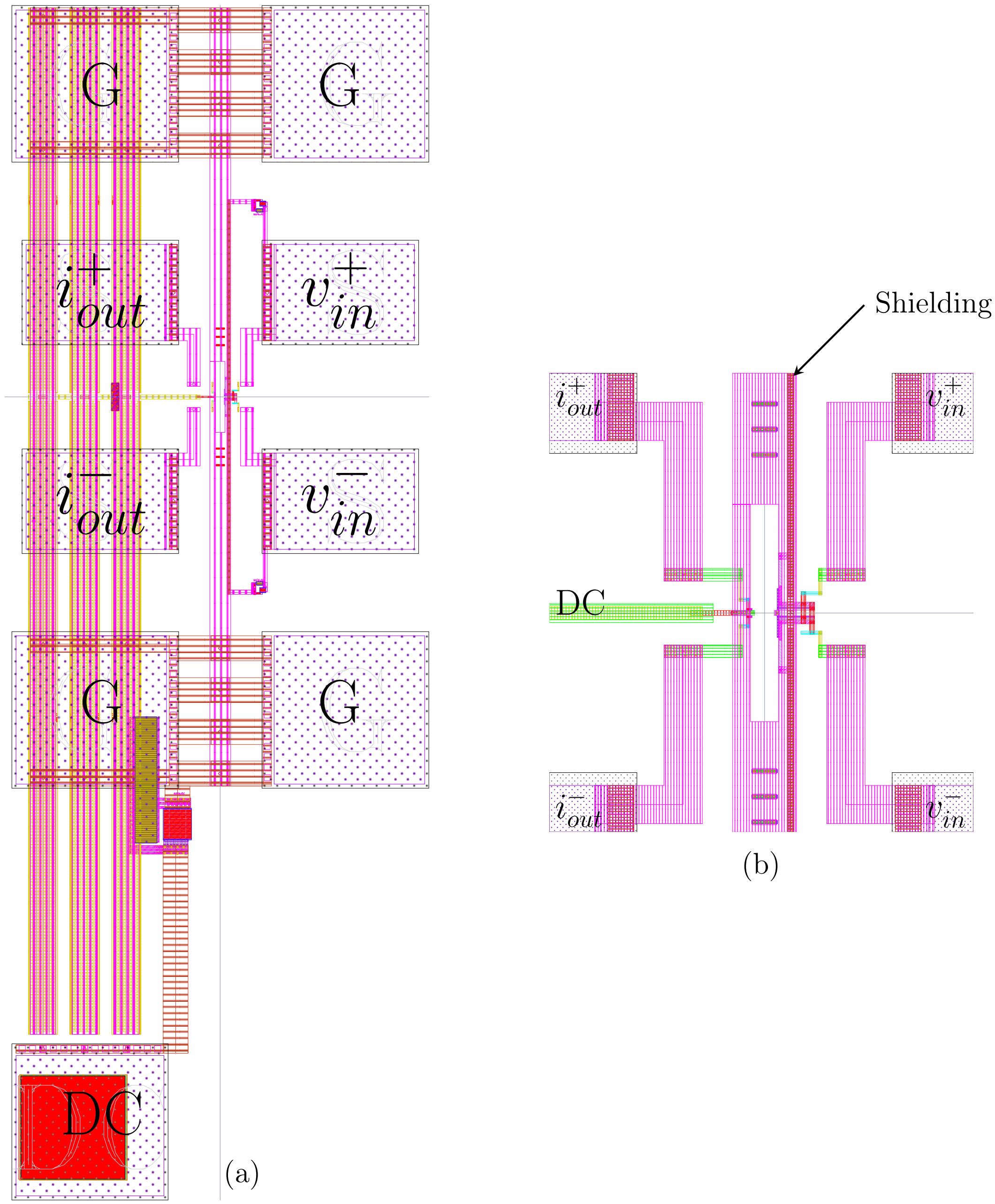


Figure 4-22: (a) Testing pads structure layout showing two sets of RF GSSG pads used for driving and sensing along with a DC pad used to provide gate bias for the sensing MOSFET. (b) Zoom-in view of the resonator area showing the different connections to the RBT as well as the shielding between input and output.

ESD protection.

On the other hand, the sensing transistor gates are the most sensitive to ESD events, due to their relatively small area. However, large on-chip capacitors are already included to reduce the noise and feed-through to this terminal. This is primarily intended for DC biasing with no RF input signals from the testing setup. ESD diodes have been added to the sensing transistor gate to ensure extra protection against ESD events.

The driving MOSCAPs are connected to the GSSG pads and are also at risk of ESD events. They represent a much larger capacitance than that of the sensing transistor gates with lower risk of ESD events. Very small ESD diodes are added to the driving MOSCAPs in order to help with ESD protection, while avoiding significantly increasing the capacitance on the driving RF port. This also helps satisfy the antenna rules of the design kit for reliable fabrication.

4.4.2 Phononic Waveguide RBTs

RBT-A and RBT-B were laid out using the P-Cell and pad structure of §4.4.1. Cross-section SEM micrographs of the fabricated RBTs in IBM 32 nm SOI are shown in Figure 4-23 and Figure 4-24. The SEM micrographs clearly show the fully differential sensing transistors in their isolated active area. The termination waveguides are also clearly highlighted for both devices.

Both RBTs occupy an area of $13.5\,\mu\text{m} \times 4.7\,\mu\text{m}$, including the fully differential driving gates routing. The width of the sensing transistors and the driving MOSCAPs is $3\,\mu\text{m}$. A fill-exclude window has been included on all layers in the resonator area to avoid any interaction with the random CMP fill.

The devices shown incorporate the regular square CMOS contacts to the active area. Devices with rectangular wall-like vias with the same dimensions have also been fabricated in the same silicon run. The rectangular vias in this IBM run demonstrated large *voids* as shown in Figure 4-25, which compromised the mechanical performance of the rectangular contacts RBTs. The major problem with these voids is that they resulted in random contact shape and density. This intern resulted in large scattering, dramatically reducing the quality factor of the resonator to the point where the output signal cannot be discerned from the noise floor.

4.4.3 RBTs with Adiabatic Transitions

RBTs with adiabatic transition have also been fabricated in IBM 32 nm SOI. A cross-section of "RBT-C" is shown in Figure 4-26. The cross-section highlights the sensing transistors as well as the termination waveguide with adiabatic taper. This implementation makes use of the CMOS regular square contacts as opposed to rectangular vias. Some contacts might appear to be missing in Figure 4-26, but this is mainly due to a small tilt in the horizontal direction in the FIB process*.

The implemented RBT occupies an area of $28\,\mu$m $\times$ $5\,\mu$m, including the fully differential driving gates routing. The resonant RBT cavity itself is $28\,\mu$m $\times$ $3\,\mu$m, the width of the sensing transistors and all MOSCAPs is $3\,\mu$m. Fill-exclude rectangles have been included on all layers to avoid any mechanical interaction with the random CMP fill.

This particular structure with adiabatic transition demonstrates the real benefit of using SKILL P-Cell with ROD for the implementation of the RBT layout: The parameters for the P-Cell are still the termination PnC metal width and separation, whereas the dimensions for the adiabatic transition sections are automatically calculated.

*This is a problem mainly because the RBT structure is very long compared to the depth of a single square via.

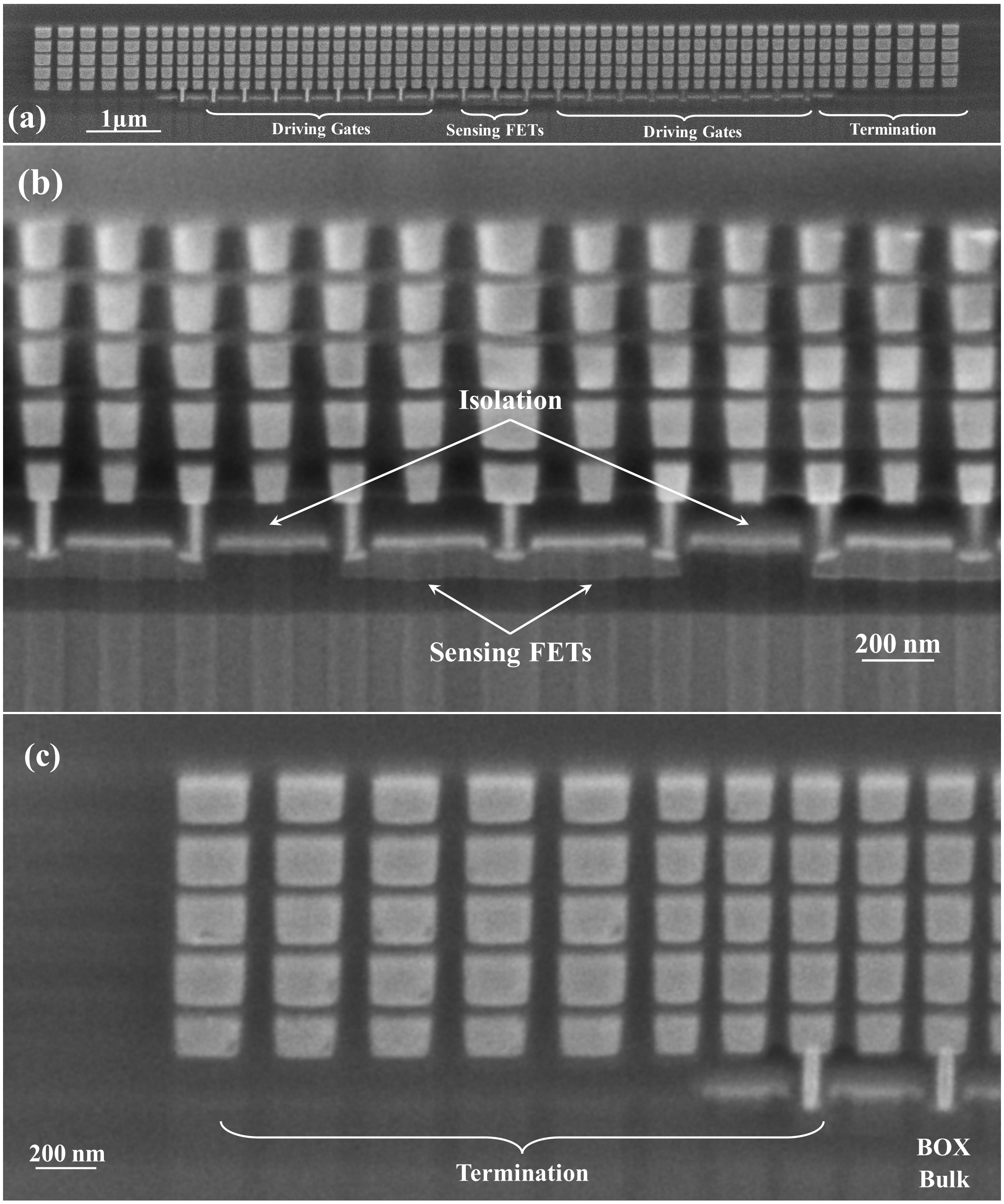


Figure 4-23: SEM micrograph for “RBT-A” showing (a) full RBT structure, (b) sensing transistors gate, and (c) termination PnC waveguide.

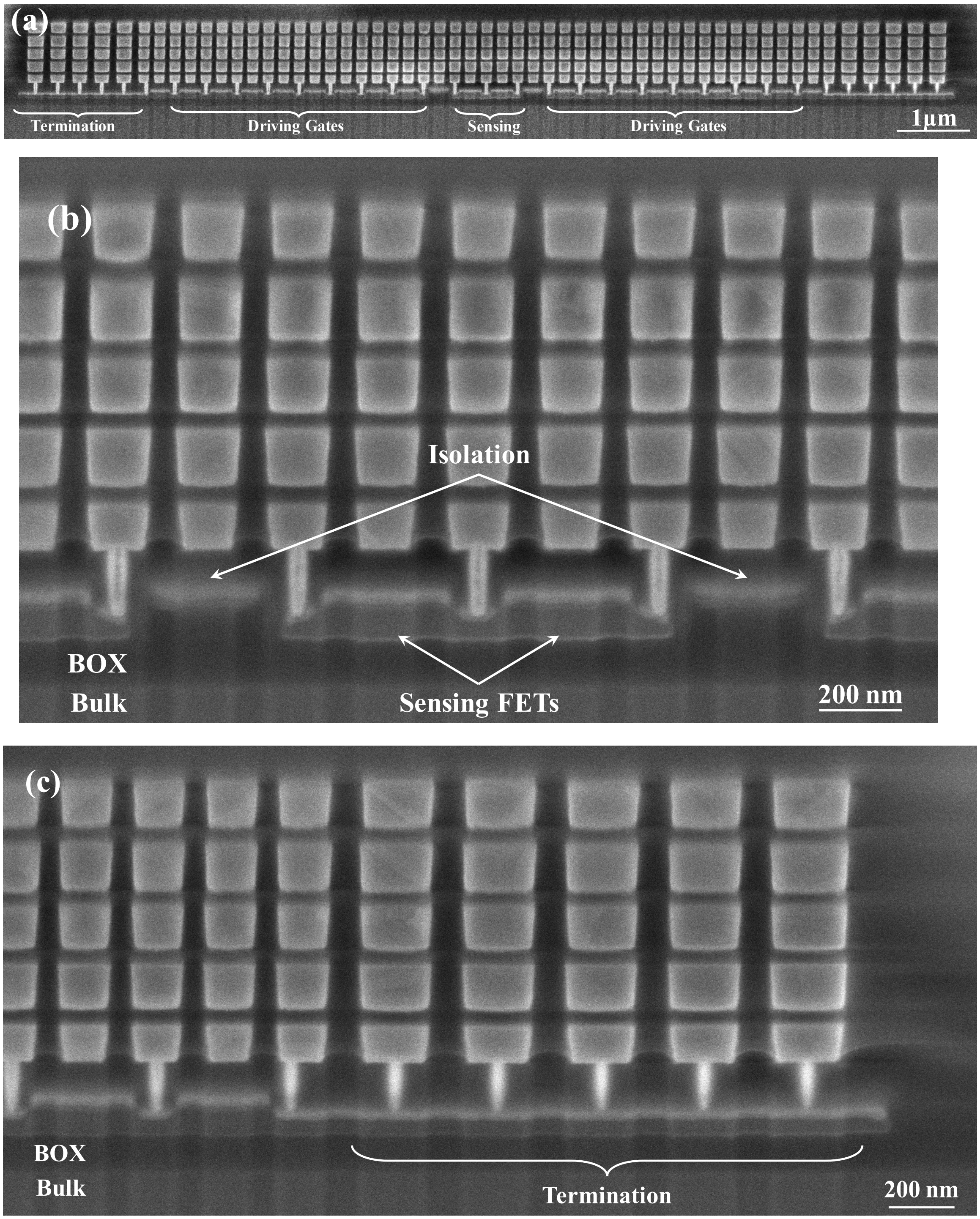


Figure 4-24: SEM micrograph for "RBT-B" showing (a) full RBT structure, (b) sensing transistors gate, and (c) termination PnC waveguide.

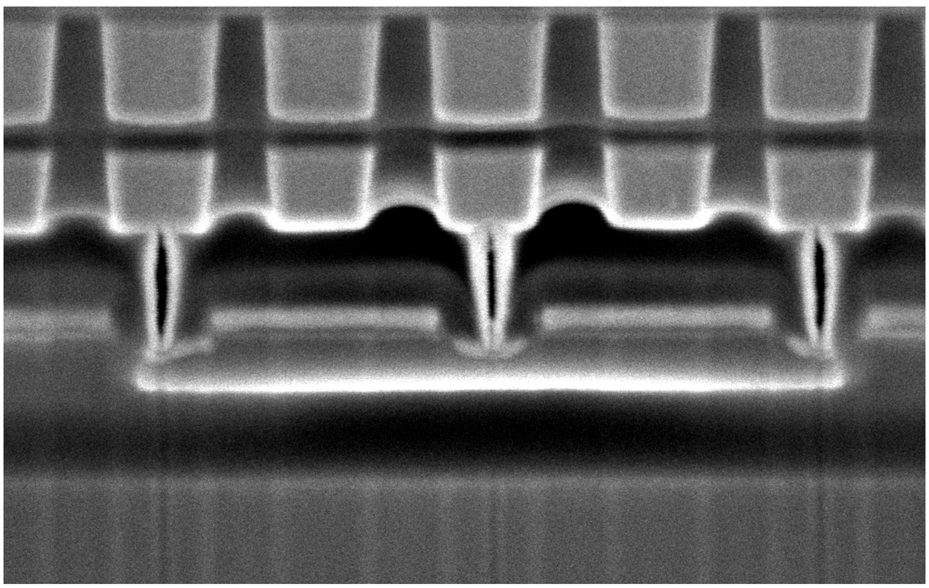


Figure 4-25: Close-up view of the sensing transistors for RBT with rectangular *wall*-like contacts. Rectangular contacts in this IBM 32 nm run ended up with large voids that compromised RBT performance.

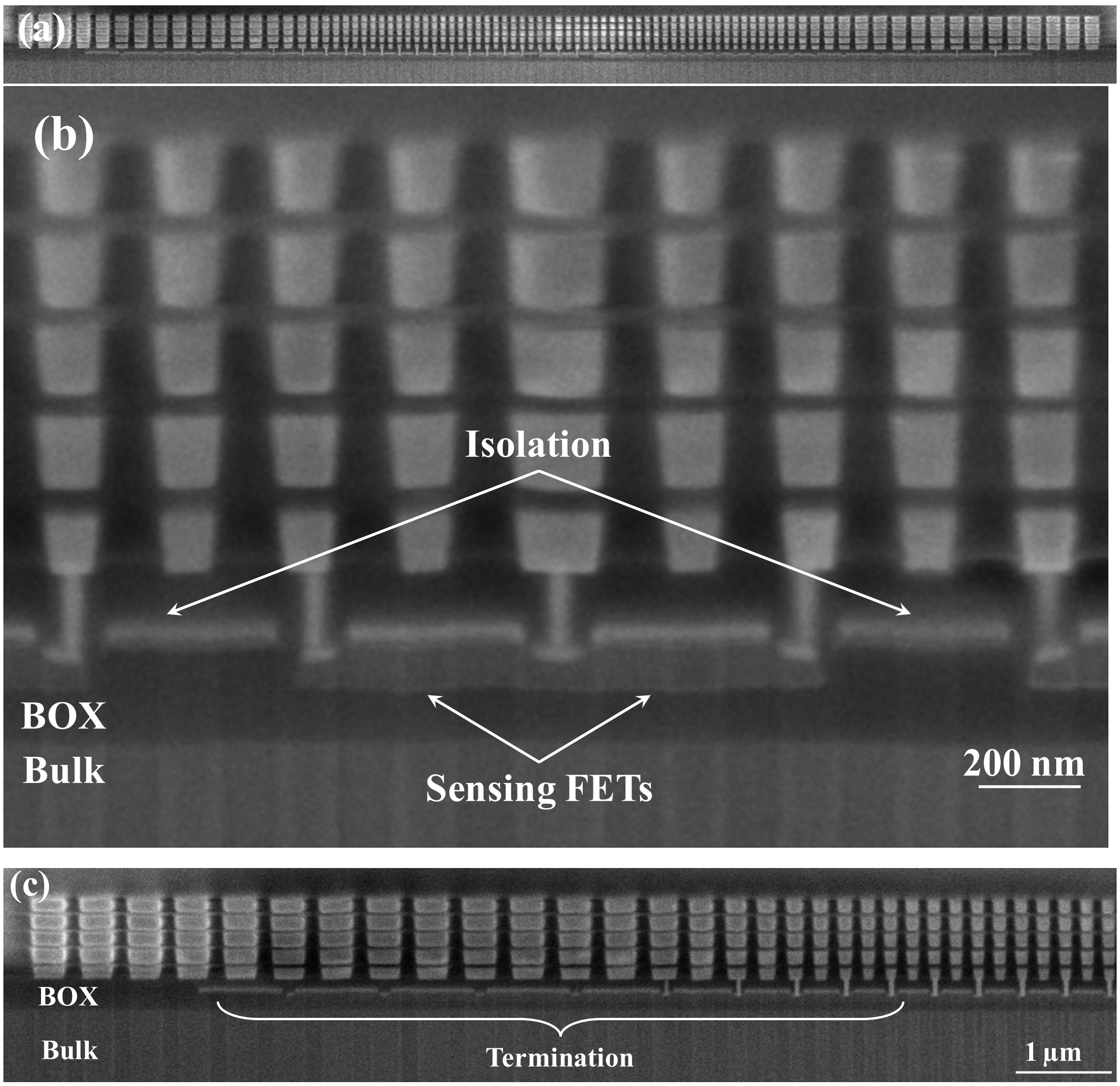


Figure 4-26: SEM micrograph for “RBT-C” showing (a) full RBT structure, (b) sensing transistors gate, and (c) termination PnC waveguide.

4.5 RBTs Characterization

RF characterization for all fabricated RBTs has been performed. This section summarizes the results as well as the testing conditions and setup.

4.5.1 DC Biasing

The CMOS RBTs considered here make use of MOSFETs for active FET sensing as well as electrostatic driving. Both sensing and driving require a DC bias to operate properly. Selecting an appropriate DC bias is the first step in the characterization of these devices.

Sensing FET Bias

Active FET sensing relies on the stress in the transistor body modulating the channel mobility through piezoresistivity. It is thus important to bias the sense transistor in a regime that maximizes the sensitivity of the drain current to the change in channel mobility. The drain current in linear and saturation regimes is given by [104]

$$I_{\text{D-lin}} = W C_{\text{ox}} \frac{v_{\text{T}}}{2(K_{\text{B}}T/q)} \left(1 - r_{\text{lin}}\right) (V_{gs} - V_{\text{T}}) V_{ds}, \tag{4.17a}$$

$$I_{\text{D-sat}} = W C_{\text{ox}} v_{\text{T}} \frac{1 - r_{\text{sat}}}{1 + r_{\text{sat}}} (V_{gs} - V_{\text{T}}) \tag{4.17b}$$

where W is the width of the transistor, C_{ox} is the gate oxide capacitance per unit area, and v_{T} is the unidirectional thermal velocity, with

$$r_{\text{lin}} = \frac{L_g}{L_g + \lambda_\circ} \quad \text{and} \quad r_{\text{sat}} = \frac{l}{l + \lambda_\circ}, \tag{4.18}$$

where L_g is the gate length, l is the length of the low field region in saturation, and $\lambda_\circ$ is the scattering mean-free-path in the channel. The sensitivity in the drain current to the mobility change is given by [104]

$$\frac{\Delta I_{\text{D}}}{I_{\text{D}}} = \frac{\Delta \mu}{\mu} \left(1 - B\right), \tag{4.19}$$

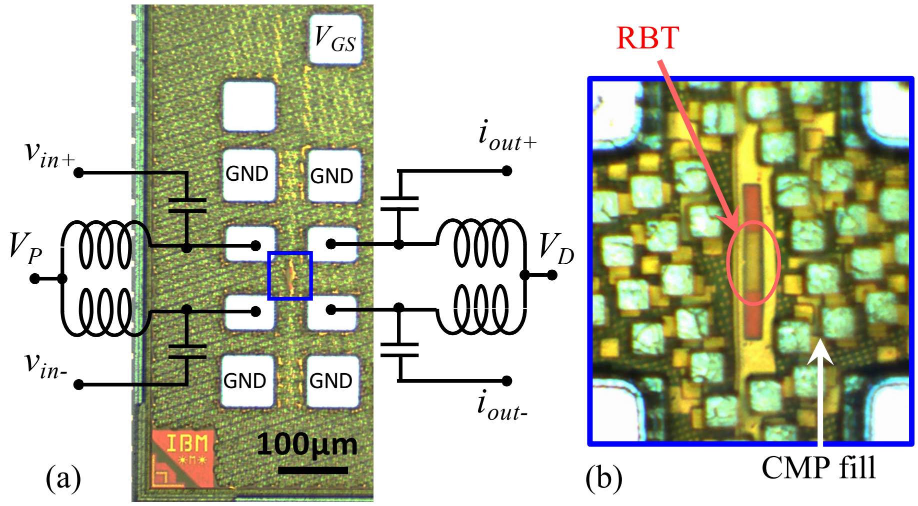


Figure 4-27: Optical micrograph showing (a) the RBT testing pads with the corresponding bias-T connections and (b) zoom-in on the RBT where the fill exclude window can be clearly seen.

where B is given by

$$B = \frac{I_{\text{D}}}{I_{\text{D-Ballistic}}} = \frac{\lambda_\circ/\zeta}{1+\lambda_\circ/\zeta}, \tag{4.20}$$

where $\zeta = L_g$ for linear regime and $\zeta = l$ for saturation.

In the linear regime, $L_g >> \lambda_\circ$ and $B_{\text{lin}} \rightarrow 0$. Thus, the sensitivity of the channel current to mobility change in linear regime becomes

$$\frac{\Delta I_{\text{D-lin}}}{I_{\text{D-lin}}} = \frac{\Delta\mu}{\mu}. \tag{4.21}$$

However, in saturation, the length of the low field region l is not generally much larger than the mean-free-path. The resulting drain current sensitivity to mobility can be smaller than that of the linear regime. Drain current sensitivity to mobility changes decreases as the device becomes quasi-ballistic and disappears in the ballistic limit $B \rightarrow 1$.

The sensing transistors channel length for the implemented RBTs is on the order of 300 nm. This suggests that the sensing transistor behavior will mostly be diffuse. However, it is still true that the drain sensitivity to the channel mobility is highest in the linear regime. For this reason, all the RBTs presented hereafter will be biased in linear regime.

The drain of the NMOS sensing transistor is selected to be biased at $V_{ds} = 85\,\text{mV}$, whereas the gate bias is usually around $V_{gs} = 700\,\text{mV}$. The gate bias is selected to yield a drain current around $100\,\mu\text{A}$ in each sensing transistor ($200\,\mu\text{A}$ in the differential pair). Meanwhile the drain bias is selected to be small enough to maintain linear operation, yet large enough to avoid existing the linear regime or turning off the transistor when the RF testing power is applied. The latter normally induce approximately $60\,\text{mV}$ peak voltage on the drain.

Drive MOSCAPs Bias

The driving MOSCAPs are also NMOS transistors and are selected to be biased in strong inversion. In strong inversion the charge (on the gate) is given by

$$Q_I(V_{gs}, x_{\text{ox}}) = \frac{\epsilon_{\text{ox}}}{x_{\text{ox}}}(V_{gs} - V_{\text{T}}), \tag{4.22}$$

where x_{ox} is the gate oxide thickness. The co-energy at a given voltage can be found to be

$$W(V_{gs}, x_{\text{ox}}) = \int_{v_{\text{T}}}^{V_{gs}} d\,V_x\; Q_I(V_x, x_{\text{ox}}) = \frac{\epsilon_{\text{ox}}}{2x_{\text{ox}}}(V_{gs} - V_{\text{T}})^2. \tag{4.23}$$

The resulting stress is found from the co-energy to be

$$T(V_{gs}) = -\frac{\partial W(V, x_{\text{ox}})}{\partial x_{\text{ox}}} = \frac{1}{2}\frac{\epsilon_{\text{ox}}}{x_{\text{ox}}^2}(V_{gs} - V_{\text{T}})^2 + \frac{\epsilon_{\text{ox}}}{x_{\text{ox}}}(V_{gs} - V_{\text{T}})\frac{\partial V_{\text{T}}}{\partial x_{\text{ox}}}. \tag{4.24}$$

Thus, the stress resulting from MOSCAPs in strong inversion is very similar to that resulting from parallel plate capacitor except that it depends on $(V_{gs} - V_{\text{T}})^2$ as opposed to V_{gs}^2. There is also an extra term resulting from the change of the threshold voltage with x_{ox}. Since V_{T} tends to generally increase with x_{ox}, the last term of (4.24) adds to the resulting stress. This addition may not be trivial and might be large enough to compensate for the lower stress due to $(V_{gs} - V_{\text{T}})^2$ as opposed to V_{gs}^2.

The small signal RF driving stress, in response to a voltage v_{in} superimposed on a DC level V_{gs}, is found by considering the first order Taylor expansion of (4.24) as

$$T_{\mathrm{ac}}(v_{\mathrm{in}}, V_{gs}, x_{\mathrm{ox}}) = v_{\mathrm{in}} \left.\frac{\partial T(V_{gs})}{\partial V_{gs}}\right|_{V_{gs}} = \frac{\epsilon_{\mathrm{ox}}}{x_{\mathrm{ox}}^2}(V_{gs} - V_{\mathrm{T}})v_{\mathrm{in}} + \frac{1}{2}\frac{\epsilon_{\mathrm{ox}}}{x_{\mathrm{ox}}}\frac{\partial V_{\mathrm{T}}}{\partial x_{\mathrm{ox}}} \times v_{\mathrm{in}}. \tag{4.25}$$

It is clear that it is still desirable to maximize the DC bias voltage V_{gs} for the MOSCAP in order to maximize the stress and hence the output signal. The maximum allowable VDD in this technology is 1.2 V. The DC bias for the driving MOSCAPs is selected to be 1 V, so that there is a reasonable safety margin for the applied RF signal.

Three independent Keithley 2400 source-measure-units (SMUs) are used to provide biasing for the sensing transistor gate, drain, and driving MOSCAPs, respectively.

4.5.2 RF Measurements Setup, Calibration and De-Embedding

The RBTs under consideration are balanced-balanced devices. However, for more reliable measurement, full 4-port S-parameters measurements have been performed for all reported devices. Calibration and de-embedding are all performed on the full 4-port S-parameters data. The de-embedded 4-port S-parameters are converted in MATLAB to fully differential 2-port S-parameters, from which a fully differential g_{em} is extracted as

$$g_{\mathrm{em}} = Y_{dd-21} - Y_{dd-12}. \tag{4.26}$$

PNA Sweep Configuration

All measurements are performed using an Agilent parametric network analyzer (PNA) model N5225A. To increase the PNA sensitivity, all the port couplers in the front panel have been reversed. Reversing the couplers increases the signal that reaches the PNA receivers by $\sim$ 15 dB, at the expense of decreasing the transmitted power from the PNA by the same amount. The signal levels to be applied to the RBTs are limited to $\sim$ 60 mV peak, corresponding to -20 dBm of transmitted power (the RBT impedance is highly mismatched). This means that the 15 dB degradation in the

PNA transmitted power can be easily compensated for by increasing the PNA power to -5 dBm in the test setup. Such a setup implies a net gain of 15 dB in sensitivity.

However, reversing the PNA couplers prohibits the use of the *integrated true mode stimulus application* (iTMSA), due to the large reflections from the reversed couplers. This makes the initial measurement for imbalance correction highly inaccurate, and measurements are thus limited to single-ended 4-port measurements.

Furthermore, in order to capture the small output signals from the RBTs, the IF bandwidth of the PNA is significantly reduced to below 100Hz. The expected high-Q of the resonators along with the uncertainty in the exact peak location force the use of large frequency sweep with a great many data points. In general, sweeps spanning as wide as 400 MHz have been used, with data points sometimes numbering up to 20,001. The large number of data points and low IF translate to very long measurement times (reaching up to 5 min for a single trace, or 20 min/device). During such long measurement times, the peak can easily shift due to temperature fluctuations. In order to mitigate this issue, a stepped point sweep is used in the PNA, which means that the PNA will measure all S-parameters for the 4 ports, before stepping to the next frequency point. This guarantees that the peak shift during measurement is minimized (unless the peak abruptly shifts to frequencies that have been measured before, which is unlikely).

It has been found that averaging tends to lower the noise floor more effectively than IF bandwidth. For this reason, an IF bandwidth of 3 kHz was always used with 30 point averages. With the stepped point sweep, the PNA measures all 4-port S-parameters 30 times before stepping to the next frequency.

PNA Calibration

The PNA is calibrated up to the tips of the GSSG probes by a full 4-port calibration procedure. A hybrid calibration algorithm developed by Cascade Microtech [105] that combines the benefits of Line-Reflect-Reflect-Match (LRRM) and Short-Open-Load-Reciprocal Through (SOLR) calibrations is used (referred to as LRRM-SOLR). Cascade Microtech WinCal software implements the LRRM-SOLR algorithm and was used for all the calibrations done in this work.

The calibration is verified on a $700\,\mu$m transmission line that is not measured during the impedance standards measurements. WinCal performs the verification by comparing the measured S_{21} and S_{34} of the transmission line to theoretical expected values. Accepted calibrations have a maximum deviation of 0.7% from theoretical expectations.

De-Embedding

Both open and short structures were included with the fabricated die for devices de-embedding. The on-chip de-embedding sites use the exact pad structure of §4.4.1, including all bypass capacitors, shielding and ESD diodes.

De-embedding is performed on the full 4-port S-parameters. The de-embedding process starts by subtracting the short impedance parameters from both the device under test (DUT) and the open, yielding

$$\tilde{\boldsymbol{Z}}_{\mathrm{DUT}} = \boldsymbol{Z}_{\mathrm{DUT}} - \boldsymbol{Z}_{\mathrm{short}} \quad \text{and} \quad \tilde{\boldsymbol{Z}}_{\mathrm{open}} = \boldsymbol{Z}_{\mathrm{open}} - \boldsymbol{Z}_{\mathrm{short}}. \tag{4.27}$$

Next, the open is de-embedded from the device as

$$\hat{\boldsymbol{Y}}_{\mathrm{DUT}} = \tilde{\boldsymbol{Y}}_{\mathrm{DUT}} - \tilde{\boldsymbol{Y}}_{\mathrm{open}}. \tag{4.28}$$

After de-embedding, the fully differential RBT electromechanical transconductance g_{em} is calculated as

$$g_{\mathrm{em}} = \hat{Y}_{21} - \hat{Y}_{12}. \tag{4.29}$$

Smoothing

To further reduce the effect of point-to-point fluctuations in measurements, a running average *smoothing* filter is applied to the measured frequency domain data. The running average filter over an odd number of samples $N = 2m + 1$, with $m \in \mathbb{N}^{+}$ has its output given by

$$y[n] = \frac{1}{N} \sum_{k=-m}^{k=m} x[n-k]. \tag{4.30}$$

In the z-domain, the filter transfer function $H(z)$ is readily found to be

$$H(z) = \frac{Y(z)}{X(z)} = \frac{1}{N} \sum_{k=-m}^{k=m} z^{-k} = \frac{1}{N} \frac{z^{m+1/2} - z^{m-1/2}}{z^{1/2} - z^{-1/2}}. \tag{4.31}$$

The filter frequency response $H(j\Omega)$ is found by setting $z = e^{j\Omega}$ as

$$H(j\Omega) = \frac{1}{N} \frac{e^{jN\Omega/2} - e^{-jN\Omega/2}}{e^{j\Omega/2} - e^{-j\Omega/2}} = \frac{1}{N} \frac{\sin\left(N\frac{\Omega}{2}\right)}{\sin\left(\frac{\Omega}{2}\right)}. \tag{4.32}$$

The frequency response of the smoothing filter is real for all frequencies, which indicates that no *delay* is introduced from input to output. This means that the location of the resonance peak will not be affected by this filter. The filter, however, may result in reduction of the resonance Qas a result of smoothing. This is not considered a problem as the reported quality factors can be treated as a conservative estimate for the actual Q of the resonators.

Figure 4-28 shows a MATLAB GUI that was designed to automate de-embedding, smoothing, and g_{em} calculation for the different RBT measurements.

4.5.3 Measured RF Response

The RBTs implemented in IBM 32 nm SOI have been characterized according to the procedure described in the previous section. Table 4.2 shows the biasing conditions for each of the RBTs under consideration.

The "RBT-A" measured g_{em} is shown in Figure 4-29. The device g_{em} with 0 V driving DC bias is considered to be the open structure and is de-embedded from g_{em} at 1 V. This de-embedding is more accurate, as a large component of the feedthrough is subtracted from the device response. Short de-embedding was found to have a negligible effect on the response. The measured g_{em} shows a resonance peak at 3.155 GHz that closely matches FEM simulation predictions. The quality factor is

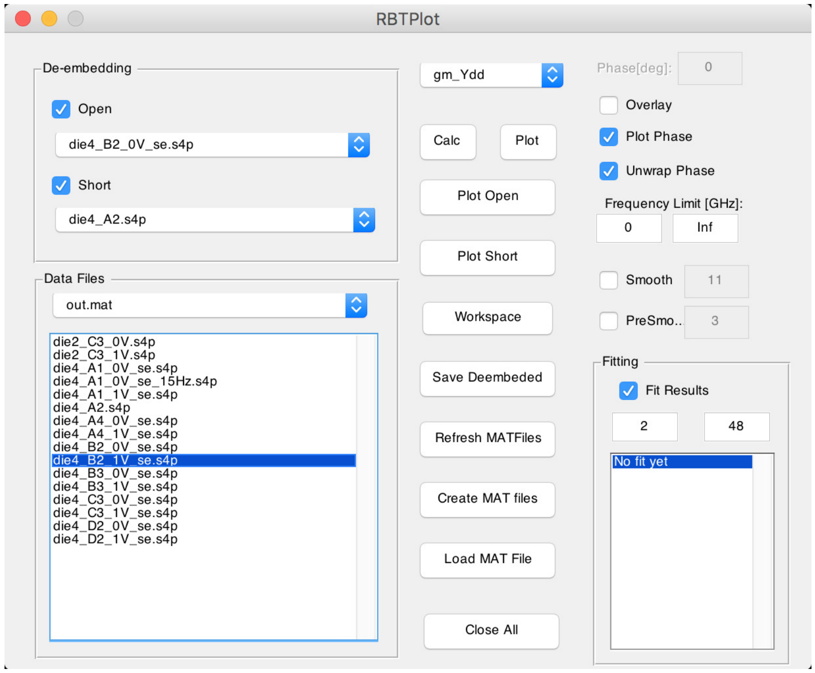


Figure 4-28: GUI used to automate de-embedding and g_{em} calculation for RBT measurement.

found to be $Q \sim 13,500$ for an $f_\circ \cdot Q \sim 4.28 \times 10^{13}$. 11-point smoothing was used for this measurement.

RBT-B has undergone similar measurements and de-embedding. The measured g_{em} is shown in Figure 4-30, where 13-point smoothing has been used. The "RBT-B" resonance frequency was found to be 3.265 GHz, closely matching FEM simulation. Quality factor $Q \sim 14,800$ is estimated from measurements for an $f_\circ \cdot Q \sim 4.85 \times 10^{13}$. "RBT-B" represents 58× quality factor improvement over the resonators of [79] and 68× improvement in $f_\circ \cdot Q$, making it the RBT with highest $f_\circ \cdot Q$ to date. This improvement clearly indicates that the quality factor and $f_\circ \cdot Q$ were not limited by material-intrinsic losses.

Finally, "RBT-C" with adiabatic transition measured g_{em} is shown in Figure 4-31. A resonance frequency of 3.089 GHz is observed, closely matching FEM simulations, whereas $Q \sim 8,950$ for $f_\circ \cdot Q \sim 2.77 \times 10^{13}$. As predicted from FEM simulations, RBT-C shows higher g_{em} than RBT-A and RBT-B due to the uniformity and stress distribution of the mode. New transduction physics (such as piezoelectric transduction by ferroelectric materials) are needed to further boost the RBTs g_{em}.

Table 4.2: DC bias for the different RBTs under consideration.

	"RBT-A"	"RBT-B"	"RBT-C"
$V_{\text{DC-drive}}$	1.0 V	1.0 V	1.0 V
V_{ds}	85 mV	85 mV	85 mV
V_{gs}	700 mV	700 mV	700 mV
I_D	231 μA	231 μA	215 μA

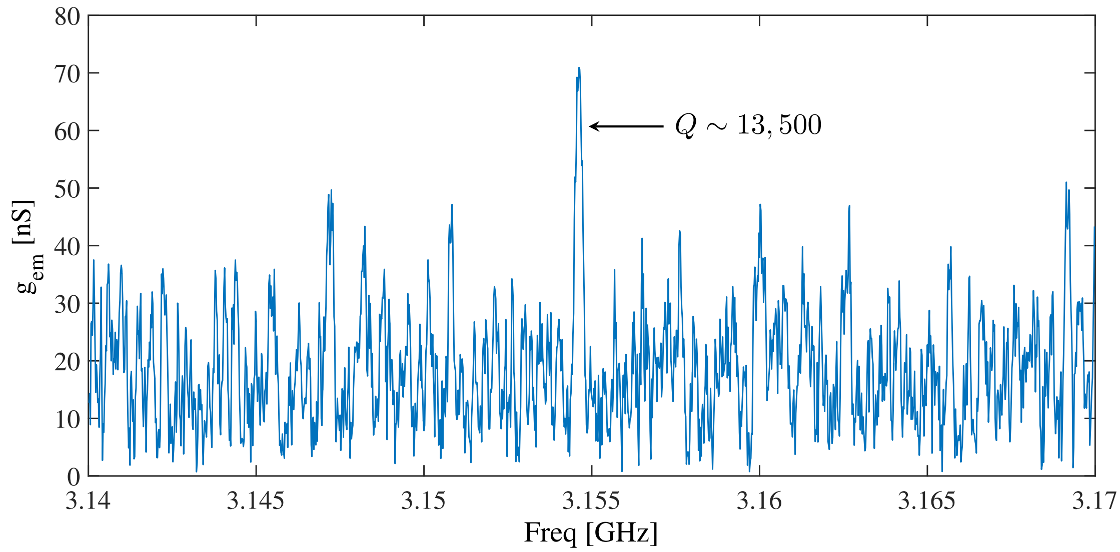


Figure 4-29: Measurement results showing fully differential g_{em} for "RBT-A," showing a resonance peak at 3.155 GHz with $Q \sim 13,500$ for $f_\circ \cdot Q \sim 4.28 \times 10^{13}$.

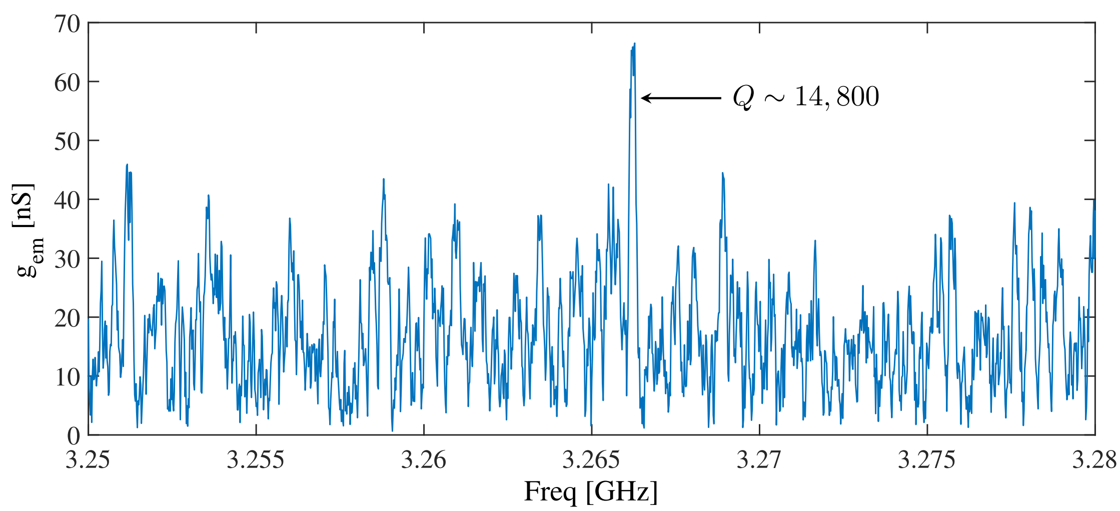


Figure 4-30: Measurement results showing fully differential g_{em} for "RBT-B," showing a resonance peak at 3.26 GHz with $Q \sim 14,800$ for $f_\circ \cdot Q \sim 4.85 \times 10^{13}$.

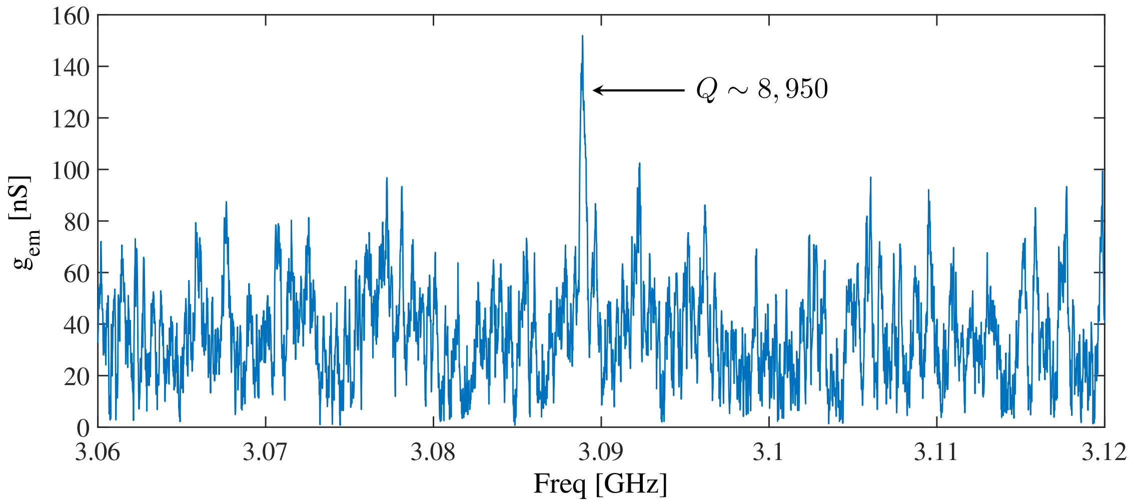


Figure 4-31: Measurement results showing fully differential g_{em} for "RBT-C," showing a resonance peak at 3.089 GHz with $Q \sim 8,950$ for $f_\circ \cdot Q \sim 2.77 \times 10^{13}$.

4.6 RBTs in FinFET CMOS Technologies

FinFET transistor technology is becoming the *de facto* standard for the 22 nm CMOS technology nodes and beyond. In these technologies, a gate is wrapped around a silicon *fin* from three directions as shown in Figure 4-32.a. This structure achieves superior electrical performance through better control of short channel effects. Unlike in planar transistors, the drain electric field is better screened by the gate, allowing tighter control of the latter on the transistor channel. This potentially reduces the subthreshold swing (S), improves the drain-induced barrier lowering (DIBL) performance, and allows for lower threshold voltage for a given off-current [106]. The fin geometry also allows for increased active area density.

In addition to the superior electrical performance, the fin structure is also capable of exceptional elastic energy confinement and waveguiding. In fact, the periodic fin structure acts as a grated waveguide, similar to those described in §3.3.3. As per the discussion in §3.4, fully-differential driving of the fin structure that enforces $\beta = \pi/a$, which ensures minimal scattering to the bulk sound cone and allows for perfect vertical confinement. Owing to the exceedingly small fin pitch in FinFET technologies, the resulting waveguide eigenmodes frequencies scale up to tens of gigahertz. Moreover, the electrostatic actuation stress is applied to the sidewalls of the individual fins.

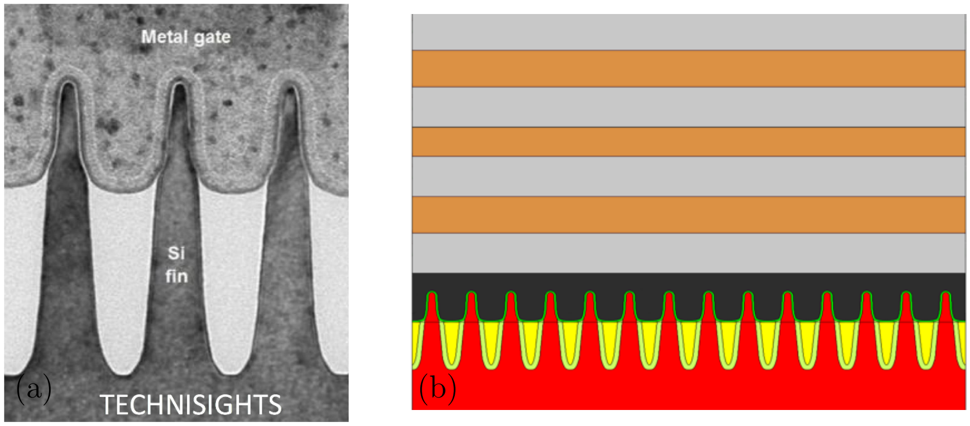


Figure 4-32: (a) Cross-section of a 14 nm FinFET CMOS technology, showing the silicon fins as well as the metal gates; and (b) Cross-section of a resonant fin transistor (RFT) in 14 nm FinFET CMOS technology.

This, together with the large active area resulting from the fin shape, allows for strong coupling to the guided modes. Such a strong coupling results in increased stresses at resonance, which yields larger signal in the sensing transistor. This directly reduces the RBT insertion loss; and subsequently, the power consumption of an oscillator based on it.

Figure 4-32.b shows the cross-section of a resonator implemented in Globalfoundries 14 nm technology. This resonator has a resonance mode at 33 GHz as seen from the FEM simulation results in Figure 4-33.a. The quality factor is limited only by the intrinsic material losses, with negligible radiation losses. The x-stress distribution at resonance is shown in Figure 4-33.b. The simulation results show that the individual fins are resonating along the FinFET *width* direction. The resonance frequency is determined by the dimensions and pitch of the transistor fin. We refer to this device as the Resonant Fin Transistor (RFT).

Moreover, RBTs similar to those of §4.3, which resonate along the gate length direction, can also be implemented in FinFET technologies. However, unlike in planar technologies, 3D FEM simulation for the fin structure is required to asses the device performance. Such a 3D FEM model is shown in Figure 4-34 for Globalfoundries 14 nm FinFET RBTs. The advantage of this design is the lithographically defined resonance frequency, which can be controlled by adjusting the periodicity of

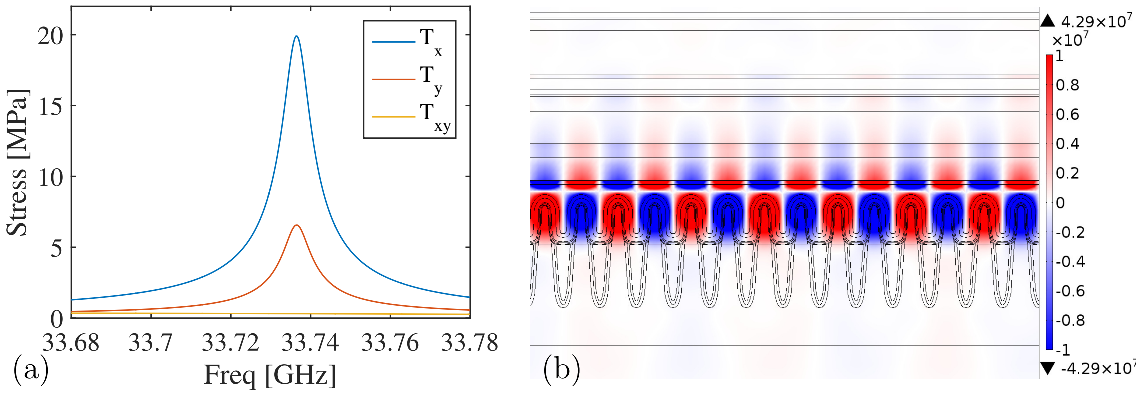


Figure 4-33: (a) FEM simulation results showing the differential average stress at sensing FETs for a 1 MPa driving stress; and (b) the x-stress for the 33 GHz resonance mode.

the transistor gates. The designed RBTs showed resonance around 10 GHz.

Multiple RFT and FinFET RBT designs have been tapedout in Globalfoundries 14 nm technology as part of a Multi-Project Wafer (MPW) run*.

* The MPW fab processing was not completed by the time of this writing.

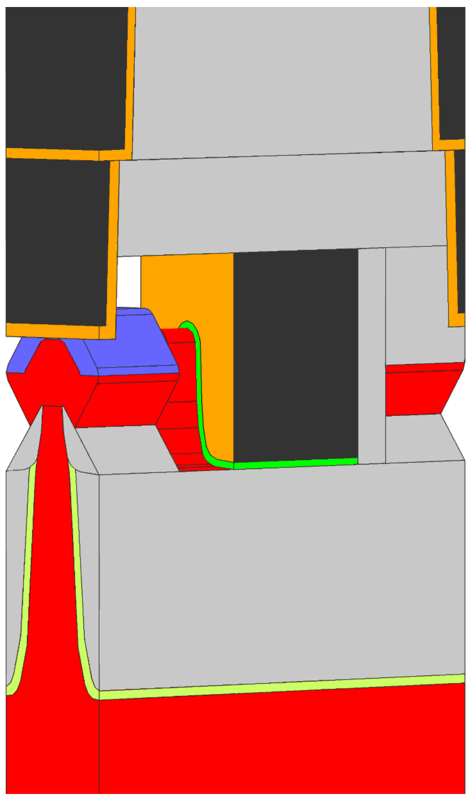


Figure 4-34: 3D FEM model used for simulating FinFET RBTs, resonating along the gate length direction.

4.7 Conclusion

In this chapter, phononic waveguides based on BEOL PnCs have been used to implement unreleased CMOS RBTs. Such waveguides confine the elastic vibrational energy vertically and form the basis for RBT cavities. Fully differential driving is proven to be necessary to excite the guided modes in these waveguides. Electrical isolation between drive and sense is also necessary; however, careful consideration is required for the structure to match the different sections of the waveguide.

Horizontal confinement can be achieved in many ways. Using another waveguide section with mismatched dispersion characteristics proves to be a good choice in terms of reducing scattering. Adiabatic transition to termination can help with scattering reduction and provides a better stress distribution for the main cavity waveguide.

Numerical optimization for CMOS RBTs is necessary to optimize the quality factor as well as the output signal amplitude. Matching the different sections of the cavity is an implicit requirement in such optimization problem. A numerical framework based on model order reduction and memoization is proposed for up to 40× speed-up in full structural optimization of the RBT. The optimization flow allows for optimizing the RBTs under consideration and can be directly applied to any MEMS device that relies on frequency domain FEM simulation for determining its performance.

Optimized RBTs have been fabricated in IBM 32 nm SOI technology without any process modifications or post-processing. RF characterization has been performed for fabricated RBTs. RBTs based on phononic waveguides show superior energy confinement. This results in fabricated devices with 58× improvement in Q and 68× in $f_\circ \cdot Q$ over previous generation RBTs. The presented CMOS RBTs have the highest Q and $f_\circ \cdot Q$ to date. It is also a solid indication that the previous generations were not limited by intrinsic material losses. The major Q improvement achieved is among the top goals of this work and paves the way for high-Q filters and low-phase noise oscillators.

Finally, RBTs in FinFET CMOS technologies have been explored. The fins structure provide superior acoustic confinement and strong coupling to the driving stress.

High Q and larger output signals are the immediate benefits. Moreover, the fins' small feature size allows for frequency scaling up to tens of gigahertz. FinFET RBTs, with resonance frequencies around 33 GHz and 10 GHz, have been implemented and tapedout in Globalfoundries 14 nm technology. Such RBTs can enable unprecedented monolithic low-phase noise, low-power oscillators at such high frequencies. These oscillators can be used for RF applications, or to build large coupled clusters for exceedingly fast unconventional signals processing.

Chapter 5

Compact Modeling of RBTs

Compact models are essential for designers who intend to use RBTs in their circuits and systems. Such models are meant to be efficient and provide insight into device performance, without the need for computationally intensive field solution. They are also required to be compatible with nodal analysis commercial spice circuit simulators, so that designers can use them seamlessly in their electronic design automation (EDA) tools. These compact models also support specialized RF and mixed signal analysis as they don't involve any hidden states. Verilog-A behavioral modeling represents an ideal environment for the development of such models. Maintaining numerical stability of these models (when solved as part of a larger system or circuit) is a critical aspect of model design.

The author previously published multiple compact models for RBTs in Verilog-A as part of the NSF NEEDS project [107–109]. The goal of this chapter is to highlight the main challenges of RBT compact models and proper approaches to addressin them. It demonstrates the transition from physics equations to nodal analysis formulation.

5.1 Modular Model

RBTs incorporate diverse and tightly coupled physics, with mechanical resonance, electrostatic actuation, piezoresistive sensing, FET operation and self-heating being

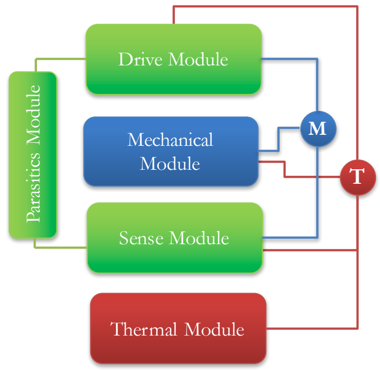


Figure 5-1: Modular model: Modules represent different physical phenomena, interacting through nodes.

the most relevant. It is also important to note that RBTs come in different structures with different driving and sensing mechanisms. For example, piezoelectric driving or sensing can be introduced, single-ended or differential driving or sensing can be used. Moreover, RBTs can be mechanically coupled or can be thermally affected by nearby power hungry circuitry. Owing to the diversity and multitude of possibilities, the RBT model has to be flexible and easily extensible. Adding new physics or changing the interaction of the existing ones should be a simple process.

For these reasons, a *modular* structure is adopted for the model. In this implementation, each of the device physics is modeled by a separate Verilog-A *module*, and the different modules are interconnected through a set of *nodes*. The *potential* of these nodes represents certain physical quantities as unknowns, whereas *branches flows* are associated with corresponding conservation laws. Such interconnection allows the different modules to interact, representing the mutual coupling between the different physics of the device.

Figure 5-1 shows the structure of a modular model in the case of an RBT. Separate modules are used to model the mechanical resonance, driving, sensing and thermal behavior of the device. The modules are connected through a mechanical displacement node and a thermal node, with the following properties:

- Mechanical displacement node:
  - **Flow**: Equivalent force/stress of the lumped device

  - **Potential**: Equivalent position of the lumped device displacement model.
  - **Conservation**: Force conservation $\sum_i F_i = 0$

- Thermal node:
  - **Flow**: Equivalent power flow in the lumped device model
  - **Potential**: Equivalent temperature in the lumped device model
  - **Conservation**: Energy conservation

Such implementation of the model allows for the seamless expansion of the model. Simple examples on such expansion possibilities include

- Adding extra drive/sense modules;
- Adding extra mechanical damping or mass loading;
- Adding extra heat sources/sinks; and
- Mechanical coupling of resonators.

The following sections of this chapter discuss the different aspects of implementing the different modules in Verilog-A.

5.2 Resonant Body Mechanical Module

The resonant body mechanical module is responsible for modeling the mechanical resonance that takes place in the body of the RBT. This is considered to be the main phenomenon for which the entire device is conceived.

5.2.1 Lumped Equivalent Circuit Models

For released RBTs, it is often possible to obtain closed form expressions for the resonance frequency and expected quality factor. This allows simple lumped equivalent circuits to be used to represent mechanical resonance. The RBT models in [107] and [108] are based on this assumption.

In these models, multiple modes are assumed to account for spurious modes and harmonics that might be useful for some non-linear applications*. Two mechanical nodes are used to model the j^{th} resonance mode as follows:

- `x[j]`: mechanical displacement node to model the amplitude of displacement of the j^{th} mode, and

- `v[j]`: mechanical velocity node to model the amplitude of velocity of the j^{th} mode.

With this separation, the equivalent model of the RBT mechanical resonance can be represented by the following differential algebraic equations (DAE)

$$\frac{d}{dt}\begin{bmatrix}\boldsymbol{v}\\ \boldsymbol{x}\end{bmatrix} = \begin{bmatrix}\boldsymbol{M}^{-1}\boldsymbol{b} & \boldsymbol{M}^{-1}\boldsymbol{K}\\ \mathbb{1} & 0\end{bmatrix}\begin{bmatrix}\boldsymbol{v}\\ \boldsymbol{x}\end{bmatrix}, \tag{5.1}$$

where $\boldsymbol{M}$, $\boldsymbol{K}$, and $\boldsymbol{b}$ are diagonal matrices representing, for each mode, equivalent mass, stiffness, and damping, respectively. The equivalent model parameters can be obtained analytically in the case of simple geometries as in [107] and [108]. In unreleased RBTs with more complex geometries, these can be obtained from fitting FEM simulations or actual device measurements. The velocity nodes are internal to the Verilog-A module† whereas the displacement nodes `x[j]` are part of the resonant body module interface, representing the mechanical nodes indicated in Figure 5-1 on page 166. Driving force will be contributed to the displacement nodes by the driving module individually, according to the respective overlap integrals. A thermal node `T` is also part of the interface of the resonant body module, to allow dynamic temperature effects on the resonance frequency and quality factor.

*By applying a signal to the sensing transistor gate, the RBT effectively becomes a mixer!

†Velocity nodes are only used to properly represent the model in DAE form. The potential of these nodes is the velocity, whereas flow is not used.

5.2.2 Transmission Line Models

The 1D longitudinal waves equations of motion in linear elastic solids can be written as [68]

$$\frac{\partial}{\partial x}(-T_{xx}) = -\rho\frac{\partial}{\partial t}v_x + F_x, \tag{5.2a}$$
$$\frac{\partial}{\partial x}v_x = -c_{11}^{-1}\frac{\partial}{\partial t}(-T_{xx}), \tag{5.2b}$$

where T_{xx} is the stress, v_x is the velocity, F_x is the body force all along the x-direction, and ρ and c_{11} are the density and stiffness matrix element of the material, respectively. Writing the equations of motion in this form allows us to see the explicit analogy to the electromagnetic transmission line telegrapher equations [68]

$$\frac{\partial}{\partial x}V = -L\frac{\partial}{\partial x}I + v_s, \tag{5.3a}$$
$$\frac{\partial}{\partial x}I = -C\frac{\partial}{\partial t}V, \tag{5.3b}$$

where V and I are the voltage and current in the transmission line, v_s is an excitation voltage, and L and C are the inductance and capacitance per unit length, respectively. By explicit comparison between (5.2) and (5.3), we get the following correspondence

$$-T_{xx} \longleftrightarrow V \tag{5.4a}$$
$$v_x \longleftrightarrow I \tag{5.4b}$$
$$F_i \longleftrightarrow v_s \tag{5.4c}$$
$$\rho \longleftrightarrow L \tag{5.4d}$$
$$c_{11}^{-1} \longleftrightarrow C \tag{5.4e}$$

This leads to the conclusion that bulk materials supporting longitudinal elastic wave propagation can be directly modeled as a transmission line with the parameters and unknowns as obtained from (5.4).

This is particularly useful in modeling 1D unreleased RBTs relying on ABRs such

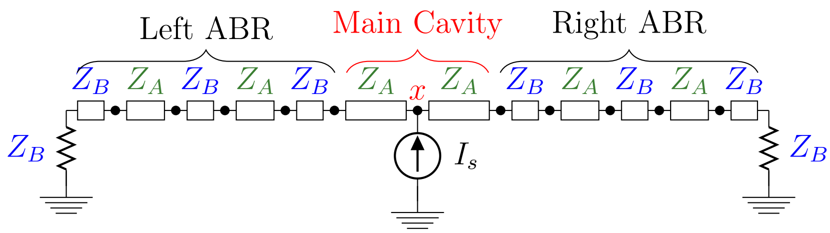


Figure 5-2: Equivalent circuit representing the waveguides used to model the mechanical resonance of the RBT in [36].

as those of [36]. The ABRs' behavior naturally follows from the transmission lines models. The 1D unreleased RBT model of [109] is based on this analogy. In this model, transmission line sections are connected by regular electrical nodes, with their potential and branch flow interpreted as stress and velocity, respectively. Transmission lines are usually available as primitives in commercial EDA simulators. Figure 5-2 shows the interconnection of the different transmission line sections, to model the RBT in [36].

Since mechanical stresses and velocity have different orders of magnitudes than voltages and currents, material properties are scaled in order to avoid tolerance and numerical problems in EDA simulators. The scaling preserves the phase velocity of the transmission lines as equal to the sound velocity in the material. On the other hand, characteristic impedance is affected by this scaling, although reflection coefficients remain the same, as the impedance scaling is uniform.

The advantage of this modeling strategy is that the model is fully parameterized in terms of the important physical device dimensions. This allows the design to study the effect of the *mismatch* of the lengths of different ABR sections as a result of process variations. This modeling technique can be extended to unreleased waveguide RBTs, assuming that the EDA simulator can support transmission lines with generic dispersion relation.

5.3 Electrostatic Transduction

The capacitive transduction module models the electrostatic internal dielectric transducer that may be used for the actuation of the RBT. The module represents the electrical interface on the input side of the device. Two implementations have been considered in [108], namely a simple parallel-plates capacitor model, and a MOSCAP model.

5.3.1 Parallel-Plates Capacitor Model

The instantaneous mechanical stress $T(t)$ induced by a parallel capacitor is given by [17]

$$T(t) = \frac{1}{2} \frac{\epsilon_{\text{ox}} v_{in}^2(t)}{g^2(t)}, \tag{5.5}$$

where $v_{in}(t)$ and $g(t)$ are the instantaneous voltage and dielectric thickness, respectively. The calculated instantaneous stress $T(t)$ is contributed back to the mechanical nodes `x[j]`, each according to its mode shape. For example, modes with odd symmetry are not affected by a driving transducer with even symmetry. In general, the component of the driving stress distribution along the resonance has to be determined through an inner product integral.

The inverse dependence on the dielectric thickness plays an important role in the model as it results in effects such as spring softening [17]. The superposition of all the mechanical mode displacements contributes to the instantaneous dielectric thickness, which is found by summing the displacement from all `x[j]` nodes as described in [107].

The module input current $i_{in}(t)$ is also calculated from $g(t)$ and the instantaneous charge $Q(t)$ as

$$i_{in}(t) = \frac{dQ(t)}{dt} \quad \text{where} \quad Q(t) = \frac{\epsilon_{\text{ox}} A}{g(t)} v_{in}(t), \tag{5.6}$$

with A being the area of the driving capacitor. The module contributes to RBT self-heating by adding power $v_{in}(t) i_{in}(t)$ to the model thermal node `T`. This power represents the total input power to the RBT's mechanical structures. The losses due to the mechanical quality factor are implicitly included in this assignment.

5.3.2 MOSCAP Drive Module Based on MVS Model

For most RBTs, electrostatic driving is implemented by means of a MOSCAP operating either in accumulation or in strong inversion. In either case, the electrostatic driving force is given by (5.5), except that the driving voltage should be interpreted as the voltage only across the gate dielectric. This requires knowledge of the MOSCAP channel surface potential. However, in most MOSFET models, in order to find the gate input current and gate capacitance, the time-dependent gate charge $Q_g(t)$ is calculated in the model [110, 111]. In this case, the electrostatic driving voltage $v_d(t)$ across the gate dielectric can be easily estimated as

$$v_d(t) = \frac{Q_g(t)}{C_{ox}}, \tag{5.7}$$

where C_{ox} is the gate dielectric capacitance. The MOSCAP dielectric thickness is also to be modulated by the strains in the RBT, yielding a time varying gate dielectric capacitance $C_{ox}(t)$

$$C_{ox}(t) = C_{ox-\circ} \frac{g_\circ}{g(t)}, \tag{5.8}$$

where $C_{ox-\circ}$ is the gate dielectric capacitance with no strain in the device. Equation (5.7) can be written more accurately as

$$v_d(t) = \frac{Q_g(t)}{C_{ox}(t)}, \tag{5.9}$$

The input current to the MOSCAP is simply found as:

$$I_{in} = \frac{dQ_g(t)}{dt}. \tag{5.10}$$

The advantage of the formulation in (5.7) and (5.8) is that unlike a channel surface potential formulation, it doesn't depend on the internal MOSFET model implementation. Instead, the gate dielectric capacitance can be found from the physical device dimensions, and the gate charge is an essential quantity for modeling gate input current. Also, the different operation regions of the MOSCAP are implicitly considered in (5.7), with the underlying MOSFET model generating the correct Q_g–V_g dependence. Moreover, $C_{ox}(t)$ may be calculated as the first thing in the MOSCAP model from (5.8), hence allowing all subsequent calculations to be dependent on the time-varying gate dielectric thickness due to the mechanical strain. Equations (5.5), (5.8) and (5.9) capture the tightly coupled physics of electrostatic driving.

An electrostatic driving model for MOSCAPs is also implemented in [108] based on the MIT virtual source (MVS) transistor model of [111]. The MVS model had to be modified to change the dielectric thickness from a constant model parameter to a dynamical unknown quantity. The MVS MOSCAP transduction module has the following nodes:

- `d,g,s,b`: the electrical terminals of the MOSCAP transducer, also the original MVS model terminals;
- `xdisp[j]`: mechanical displacement node for the j^{th} harmonic mode as described earlier;
- `T`: Thermal node.

All of these nodes are available at the interface of the MVS MOSCAP drive module. The internal workings of this module can be summarized as follows:

1. Update the dimensions and parameters according to the instantaneous temperature value.
2. Find the instantaneous dielectric thickness as in [107] and [108].
3. Find the instantaneous gate dielectric capacitance based on (5.8).
4. Go through all the MVS model calculations, concluding by finding the gate charge.
5. Find the driving voltage based on (5.9).
6. Calculate the electrostatic force and apply it to all the mechanical nodes `xdisp[j]`, with the correct weighting factors.

7. Calculate the instantaneous power as $v_{in}(t)i_{in}(t)$ and supply it as a power contribution to the device thermal node T.

The input current is already calculated and assigned to the gate terminal (`g`) as part of the MVS model calculations.

5.3.3 Numerical Considerations: Inverse Dielectric Thickness Dependence

Another major challenge for the implementation of this module is that the force term in (5.5) includes the inverse square of the instantaneous thickness $g(t)$, which is a model unknown to be solved for. There are three major issues with such dependence:

1. The square of the thickness $g^2(t)$ makes both positive and negative thicknesses *mathematically* valid solutions. In this case, the simulator may converge to non-physical solutions incorporating negative dielectric thickness.

2. Singularity at $g(t) = 0$: If during the process of finding a solution the simulator plugs-in the wrong numbers, a division by zero may result.

3. The inverse square of the thickness $g^2(t)$ may also generate very large numbers, causing the nodal matrices to become ill-conditioned.

To address the first two issues, a smoothing function is used for the dielectric thickness with the form

$$\mathrm{smoothg}(g) = \frac{g + \sqrt{g^2 + c^2}}{2}. \tag{5.11}$$

This smoothing function is plotted in Figure 5-3. To address the risk of generating large numbers, the natural logarithm together with the Verilog-A `limexp` function are used:

$$\frac{1}{g^2(t)} \Longrightarrow \texttt{limexp}\left(-2 \times \ln g(t)\right). \tag{5.12}$$

Both the smoothing function of (5.11) and the limiting technique of (5.12) are also used for the evaluation of the instantaneous gate dielectric thickness for the MVS MOSCAP driving model in (5.8).

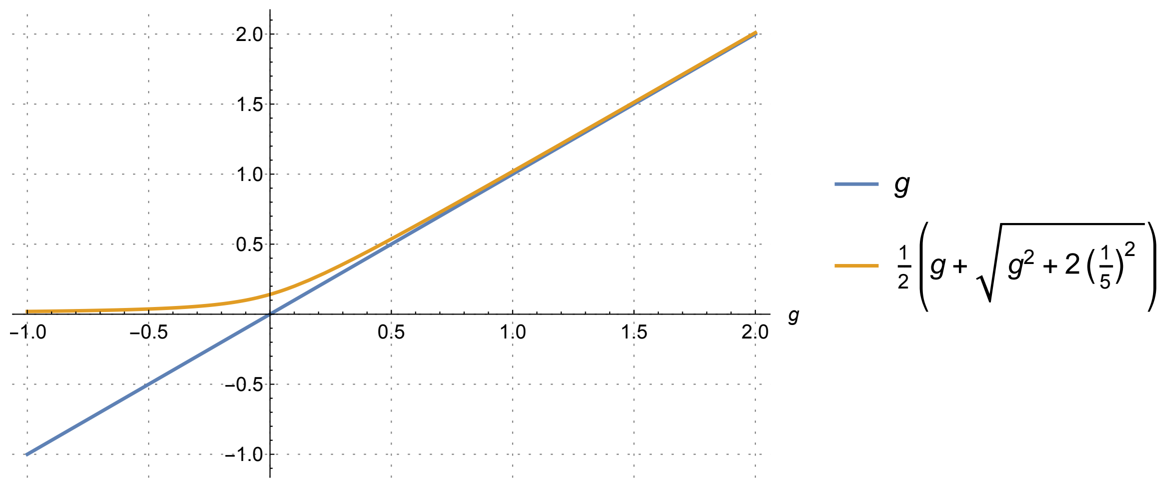


Figure 5-3: Smoothing function used to break the symmetry of $1/g^2(t)$ and also removes the singularity at $g(t) = 0$.

5.4 Active FET Sensing

FET sensing relies on having the mechanical vibrations in the cavity modulate the carrier mobility in the sensing transistor channel through piezoresistivity. As pointed out in §4.5.1, when the sensing FET is biased in linear regime, drain current modulation is proportional to mobility variations as in (4.21). However, the exact dependence of the drain current on channel mobility is a strong function of the biasing conditions. For this reason, a modified version of the BSIM model [110] as well as the MVS model [111] have been used to model the sensing FET as in [107, 108]. Using the BSIM model is useful when considering CMOS RBTs, since most foundries provide BSIM model cards for their MOSFETs. This greatly simplifies the task of parameter extraction for the RBT model.

Piezoresistive mobility modulation depends on the relative direction of the stress and current flow. For the RBT structures under consideration, the current flow along the FET channel is aligned with the T_{xx} stress. On the other hand, T_{yy} stress represents the transverse direction. The local mobility modulation in the channel is found to be

$$\frac{\Delta\mu(x)}{\mu} = -\pi_{\text{long}}\, T_{xx}(x) - \pi_{\text{trans}}\, T_{yy}(x), \tag{5.13}$$

where π_{long} and π_{trans} are the longitudinal and transverse piezoresistive coefficients.

These are found according to coordinate transformations described in appendix B.

In the case of 1D RBTs as in [107,108], where the mechanical nodes `x[j]` represent the displacement in x-direction, the average stress in the FET channel T_{ch} can be approximated as

$$
\begin{aligned}
T_{\text{ch}} &= \frac{1}{L_{\text{FET}}} \int_0^{L_{\text{FET}}} dx\; T_{xx}(x) && \text{(5.14a)} \\
&= \frac{Y}{L_{\text{FET}}} \int_0^{L_{\text{FET}}} dx\; \frac{\partial u(x,t)}{\partial x} && \text{(5.14b)} \\
&= \frac{Y}{L_{\text{FET}}} (u(L_{\text{FET}}, t) - u(0,t)) && \text{(5.14c)}
\end{aligned}
$$

where L_{FET} is the FET channel length and $u(x,t)$ is the instantaneous x displacement. The latter can be found from the mechanical nodes `x[j]` by considering a specific mode shape. On the other hand, the mechanical node `x` for the implementation in [109] directly represents the stress in the FET channel. In this case, only a proportionality factor is required to estimate the average stress from the potential of `x`.

The FET sensing module has the following nodes available at its interface:

- `d,g,s,b`: the electrical terminals of the sensing FET, also the original MVS model terminals;
- `x`: equivalent mechanical displacement node(s);
- `T`: Thermal node.

The module operation can be summarized as follows:

1. Update the dimensions and material properties according to the instantaneous temperature value.
2. Find the instantaneous channel mobility based on (5.13).
3. Evaluate the full original BSIM model or MVS model code with the instantaneous mobility.

Smoothing and limiting functions are also used to protect the simulator from numerical hazards [108, 109].

5.5 Thermal Model and Electrical Parasitics

Temperature and self-heating are important physics to consider for RBT behavioral modeling. Change of temperature affects the sensing FET characteristics as well as the resonance frequency, due to change in both the resonator dimensions and material properties. Accounting for thermal effects of the RBT would basically involve solving the heat diffusion equation [17]

$$\frac{\partial T}{\partial t} = \frac{\kappa}{\widetilde{C}}\nabla^2 T + \frac{1}{\widetilde{C}}\widetilde{P}, \tag{5.15}$$

where T, κ, $\widetilde{C}$ and $\widetilde{P}$ represent the temperature, thermal conductivity, heat capacity, and volumetric power sources.

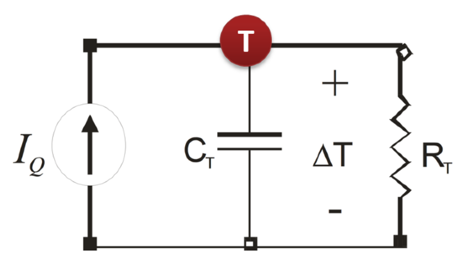


Figure 5-4: Thermal equivalent circuit for the RBT.

The thermal diffusion equation can be easily lumped into an R-C network as that of Figure 5-4, where temperature is modeled as a voltage, and thermal conductivity and heat capacity as an equivalent resistance and capacitance, respectively. The equivalent circuit current represents the power [17]. This modeling strategy is widely used, and both the BSIM model and MVS model rely on it to model self-heating effects. The thermal node is shared between the different RBT model modules along with MVS and BSIM models.

The temperature obtained from the equivalent circuit is used to assess the self-heating effects on the different model components. For the different physical parameters of the model and material properties, the following relation is assumed:

$$X = X_\circ\big(1 + \alpha_{X-1}\Delta T + \alpha_{X-2}(\Delta T)^2 + \alpha_{X-3}(\Delta T)^3 + \cdots\big), \qquad (5.16)$$

where X represents the effective value of the physical parameter including temperature effects, $X_\circ$ the value of the physical parameter at nominal temperature, and α_{X-i} represents the thermal coefficients of the physical parameter under consideration. This relation is applied for dimensions, density, and elastic properties as in [107, 108].

Finally, electrical parasitics are considered. Feed-through from the RBT input to the sensing FET gate, drain and body are particularly important. The RBT compact model includes a complete network of parasitic feed-through as that shown in Figure 4-6. This parasitics network is modeled as a standalone module that connects to the drive and sense module and implements the various capacitors between the respective nodes.

Example simulation results from the model implemented in [109] are shown in Figure 5-5 and Figure 5-6.

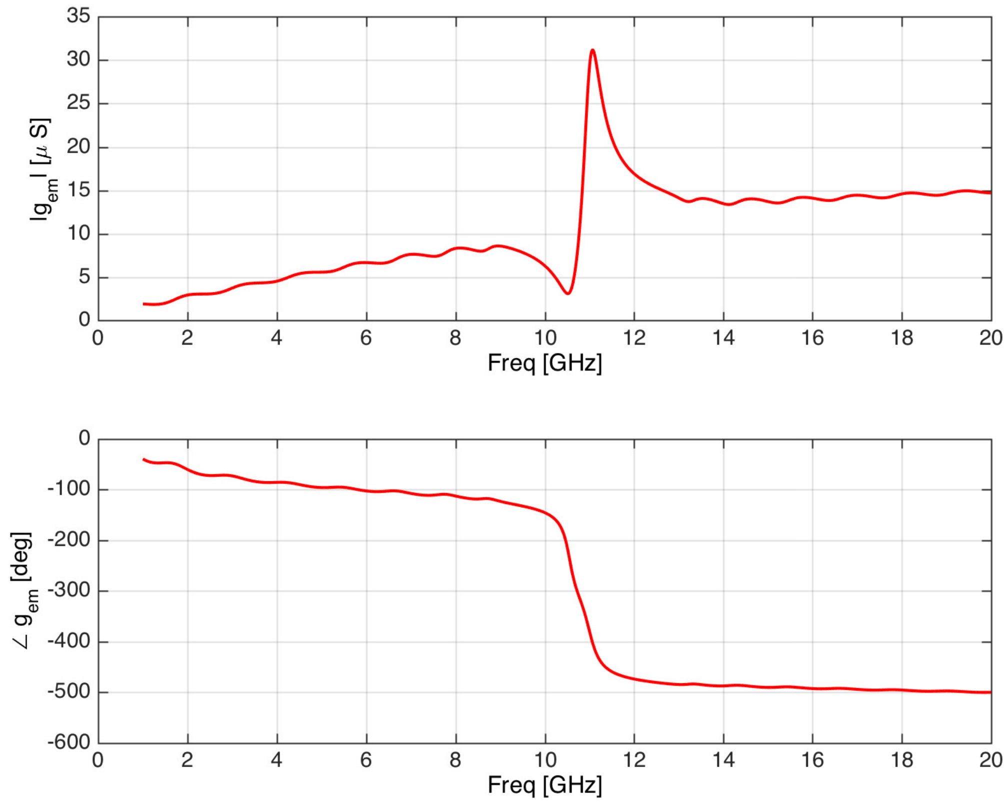


Figure 5-5: Magnitude and phase of the RBT transconductance g_{em} with default model parameters as in [109].

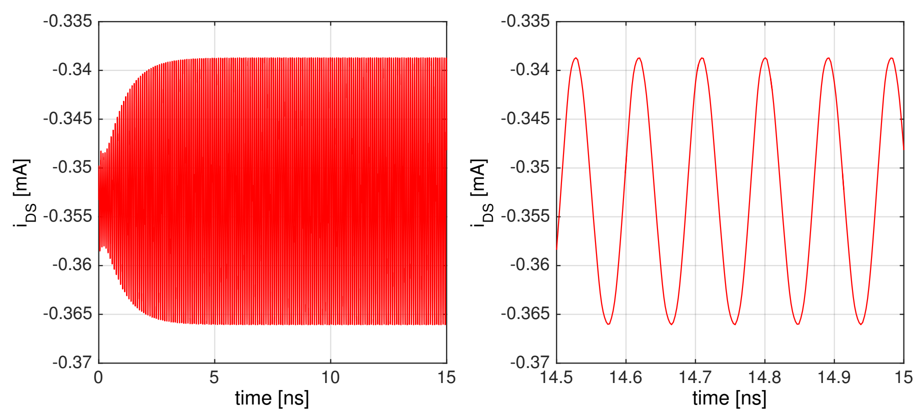


Figure 5-6: Transient drain current in response to driving the RBT with $500\,mV$ sinusoidal signal at its resonance frequency as in [109].

5.6 Conclusion

A practical approach towards RBT compact modeling has been presented. Such models allow fast and efficient simulation of the tightly coupled physics characteristic of RBTs, without the need for computationally intensive FEM field solutions. The presented models are cast into a graph formulation with nodes and branches that is fully compatible with commercial EDA software. They also support specialized RF and mixed signal analysis, as they don't involve any hidden states.

A modular approach has been presented, where individual physics are modeled by separate modules, interacting through a set of nodes. Such implementation allows for model expansion, incorporation of new physics, and coupling multiple RBT instances.

Mechanical resonance has been modeled either as a mechanical equivalent circuit or based on transmission lines. Equivalent circuit models work best for simple geometries or for fitting measured data, whereas transmission line models provide more accurate parameterization of the model and allow for fast estimation of mismatch and process variations.

Models for electrostatic transduction have been presented with special attention to the numerical integrity of the model. Both parallel-plate capacitor and MOSCAP transducers have been considered. Next, modeling FET sensing has been discussed. This relies on the full BSIM or MVS transistor model, while changing the mobility from a constant model parameter to a simulation unknown to be determined based on the instantaneous stress in the FET channel. Finally, an equivalent circuit model has been adopted for self-heating, while a full parasitics network is also included.

The presented compact models allow circuits and systems designers to seamlessly integrate them into their workflow, providing a fast and reliable solution to evaluate the performance of circuits and entire systems.

Chapter 6

MEMS-Based Oscillators and Filters in GaN MMIC Technology

Gallium nitride (GaN) is becoming an increasingly popular material for RF monolithic microwave ICs (MMICs) and power electronics. With the ever-growing demand for systems with higher efficiency and smaller foot-print, monolithic GaN high-Q filters and low phase-noise oscillators are highly desirable. Being a piezoelectric material with favorable electromechanical properties, GaN lends itself as a potential material for high-Q MEMS resonators [43]. While multiple MEMS-based oscillators have been demonstrated [112–117], this chapter presents the design and prototyping of the first monolithic GaN MEMS-based MMIC oscillator and filters for RF applications [52].

6.1 MEMS in GaN MMICs

Gallium nitride's electrical properties make it excel in both MMICs and power electronics. More specifically, the AlGaN/GaN heterostructure naturally forms a 2D electron gas (2DEG) that is confined in a potential well at the GaN surface [118]. The lack of ionized impurity scattering yields high mobility for electrons in the 2DEG layer, making it ideal for the channel of high electron-mobility transistors (HEMTs) [119]. Together with high 2DEG electron sheet charge concentration and saturation velocity, GaN HEMTs demonstrate superior performance in microwave applications. Also as a

Table 6.1: Electrical properties of wurtzite GaN compared to Si and GaAs

	Si	GaAs	GaN
Channel e- Mobility ($cm^2/V{\cdot}s$)	350	10^4	2000^*
Saturation velocity (cm/s)	10^7	1.5×10^7	2.5×10^7
Sheet charge concentration (cm^{-2})	10^{13}	10^{12}	10^{13}
Bandgap (eV)	1.12	1.43	3.4
Breakdown electric field (MV/cm)	0.3	0.4	3.3

Table 6.2: Electromechanical properties of wurtzite GaN compared to Si and GaAs

	Si	GaAs	GaN
Density (kg/cm^3)	2320	5320	6150
Elastic Modulus c_{33} (GPa)	398	118	165
Acoustic velocity (m/s)	8415	2470	8044
Piezoelectric coefficients e_{33} (cm^{-2})	-	-0.16	0.65
Transduction efficiency k_{eff}^2 (%)	-	0.04	2

large bandgap material, it can withstand high electric fields before breakdown, which when coupled with the high ON-current makes it ideal for power electronics. Table 6.1 summarizes the important electrical properties of GaN as compared to silicon and GaAs.

Moreover, the wurtzite crystal structure of GaN makes it a piezoelectric material with good electromechanical transduction efficiency ($k_{eff}^2 \sim 2\%$). Combined with its high acoustic velocity ($8000\,m/s$), large elastic modulus ($c_{33} \sim 400\,\mathrm{GPa}$), and low intrinsic acoustic losses, GaN is a potential candidate for high-Q MEMS resonators [43]. Table 6.2 compares GaN electromechanical properties to those of Si and GaAs. In addition to the intrinsic GaN properties, the 2DEG layer of the AlGaN/GaN heterostructure was demonstrated to be beneficial for switchable piezoelectric transduction [47], paving the way for reconfigurable filters.

The unique properties of GaN as a wide bandgap semiconductor and a good piezoelectric material streamlines the monolithic integration of MEMS resonators and inertial sensors with RF circuits. Multiple groups have demonstrated MEMS resonators in GaN MMIC technology, where the 2DEG of GaN HEMTs were even used as transducers for driving and sensing acoustic vibrations [13, 44–51].

*with AlN interfacial layer

A gold-free (Au-free) GaN MMIC-MEMS process was developed at MIT Microsystems Technology Laboratories (MTL) by L.C. Popa [44]. The fabrication process is summarized in Figure 6-1. It starts with a shallow AlGaN mesa etch defining the HEMTs active area, followed by ohmic and gate metal deposition and patterning. The die surface is passivated, followed by a deep GaN etch to define the MEMS structure. A final XeF_2 isotropic silicon etch is performed to release the MEMS resonators. This process is able to integrate MEMS device with minimal modification to the regular GaN MMIC HEMT flow. All devices and circuits presented in this chapter were fabricated at MIT MTL by L.C. Popa in the aforementioned process.

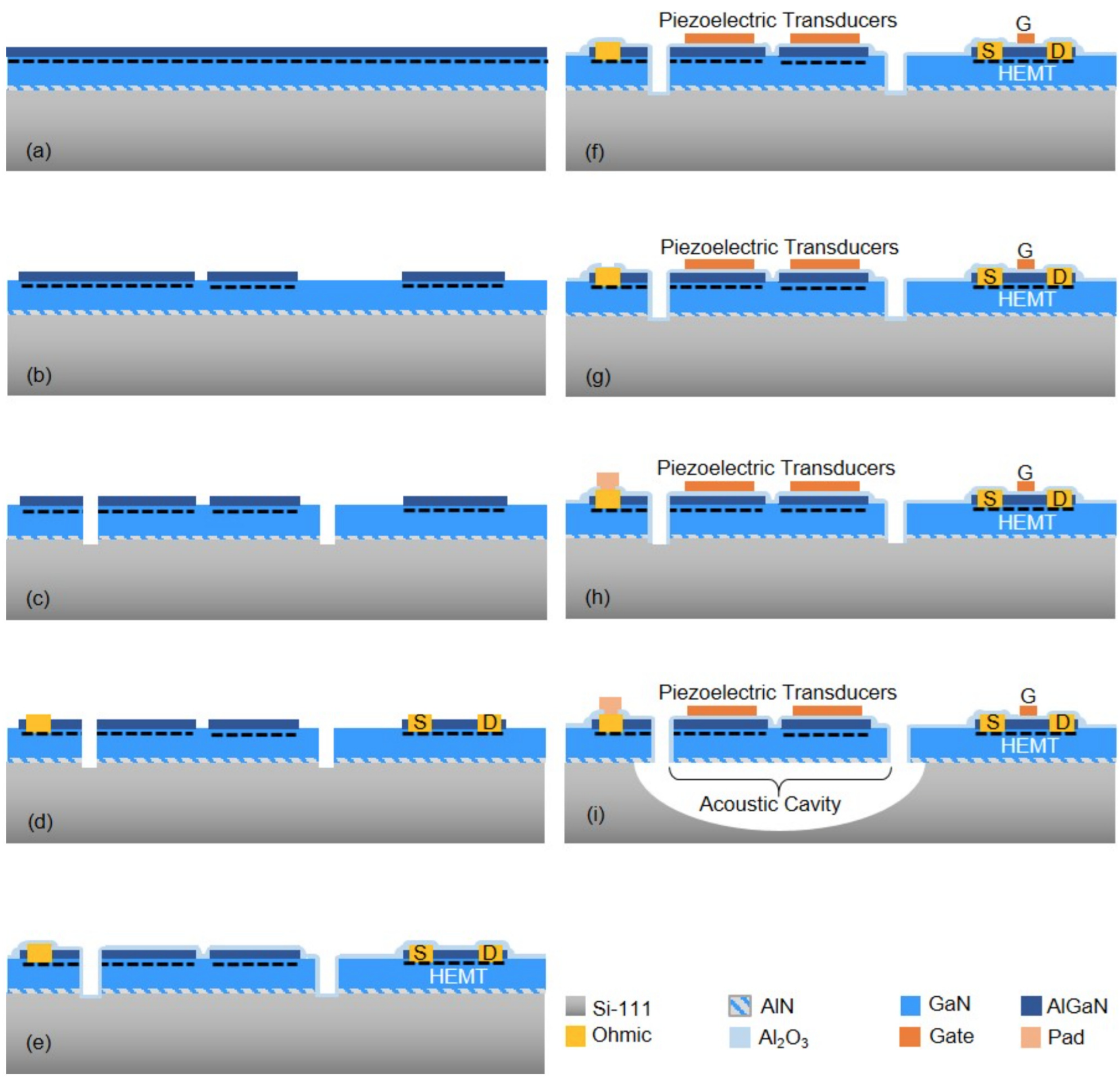


Figure 6-1: Au-free GaN MMIC-MEMS fabrication process developed by L.C. Popa at MIT MTL after [44].

6.2 Lamb-mode Resonators

Lamb-mode resonators recently gained lots of attention, with AlN and GaN being common choices for resonator material. In contract to thickness mode resonators, Lamb-mode resonators have their resonance frequency defined by horizontal dimensions. This allows for multiple frequencies on the same chip at virtually no-cost for muliband operation.

6.2.1 S_0 Lamb-Mode Resonator

The symmetric Lamb-mode (S_0) is generally favored over the antisymmetric Lamb-mode (A_0) [44, 120, 121]:

- The S_0 mode has higher phase velocity, which enables scaling to high frequencies while maintaining reasonable dimensions for fabrication.

- The S_0 mode shows lower phase velocity dispersion for normalized GaN thickness (h_{GaN}/λ), which minimizes resonance frequency dependence on the GaN thickness and hence lower sensitivity to process variations.

Additionally, the S_0 mode has high-Q and moderate electromechanical transconductance efficiency k_{eff}^2. The elimination of gold from the MMIC process reduces the acoustic losses further enhancing $f_\circ \cdot Q$ of MEMS resonators as demonstrated with record breaking performance devices in [44]. Figure 6-2 shows two variants of Lamb-mode resonators where a bottom electrode may be included. Although such electrode improves the transduction efficiency k_{eff}^2 [44, 121], it is omitted from the resonators used in this work, leading to *top-drive only* resonators. This is intended to maintain maximum compatibility with existing GaN MMIC technologies by avoiding modifications to the GaN stack and reducing process complexity. An SEM of the full Lamb-mode resonator is shown in Figure 6-4 [44].

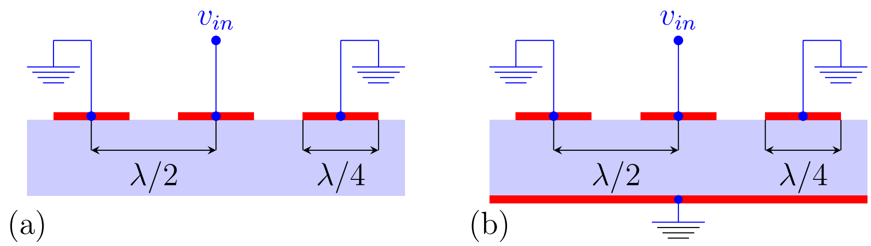


Figure 6-2: Lamb-mode resonators (a) top-drive only resonator (b) grounded bottom electrode included.

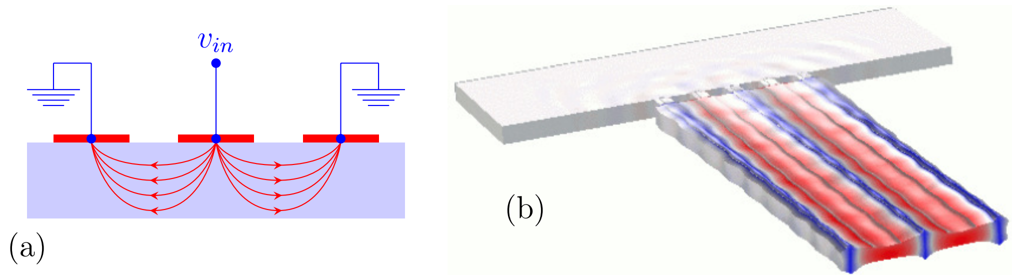


Figure 6-3: Top-drive only Lamb-mode resontor (a) Electric field lines (b) S_0 mode stress distribution.

6.2.2 Equivalent Circuit Model

With IIP_3 exceeding +20 dBm [44], the aforementioned Lamb-mode resonators are sufficiently linear for the purpose of the low-power oscillators design under consideration. Figure 6-5 shows a linear modified Butterworth-Van-Dyke (MBVD) equivalent circuit model that is sufficient to model the resonator for all practical oscillator design purposes. Full 2-port S-parameters RF characterization was performed for the resonators and fitted to the MBVD equivalent circuit. Figure 6-6 shows the measured admittance Y_{21} of one of the Lamb-mode resonators considered, as compared to the fitting results from the MBVD circuit. This particular resonator has a resonance frequency of 1.019 GHz with a quality factor $Q \sim 4200$. The complete parameters for the MBVD equivalent circuit are listed in Table 6.3.

The dominant equivalent circuit parameters scale to a first order with the non-resonant dimension w and the number of periods N of the Lamb-mode resonator, as

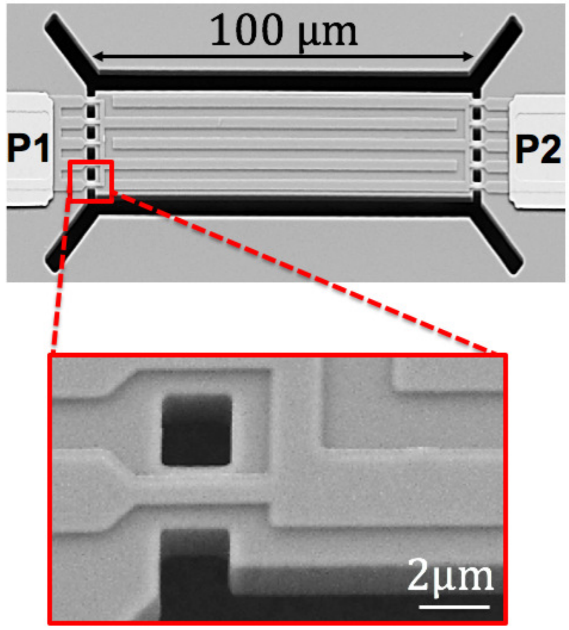


Figure 6-4: Lamb-mode resonator SEM after [44]

follows:

$$R_m(N,w) = R_{m\circ}\frac{N_\circ w_\circ}{Nw} \tag{6.1a}$$
$$C_m(N,w) = C_{m\circ}\frac{Nw}{N_\circ w_\circ} \tag{6.1b}$$
$$L_m(N,w) = L_{m\circ}\frac{N_\circ w_\circ}{Nw} \tag{6.1c}$$
$$R_f(N,w) = R_{f\circ}\frac{N_\circ w_\circ}{Nw} \tag{6.1d}$$
$$C_f(N,w) = C_{f\circ}\frac{Nw}{N_\circ w_\circ}, \tag{6.1e}$$

where the subscript $\circ$ corresponds to a given resonator design. Both w and N are lithographically defined dimensions and can be set by the designer. This adds a design degree of freedom as the resonator complex impedance can be tuned (favoring smaller R_m or smaller C_m), allowing for MEMS-circuit co-design and co-optimization.

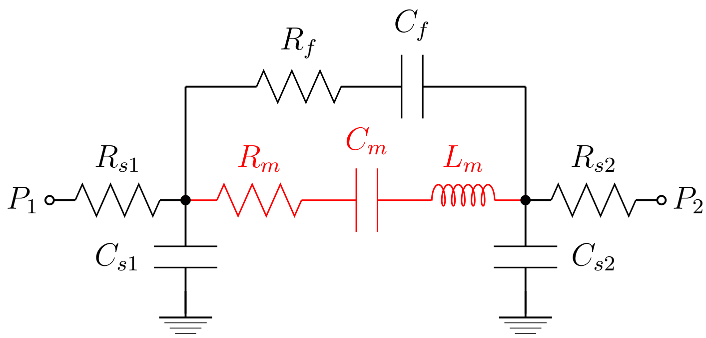


Figure 6-5: Equivalent circuit for Lamb-mode resonator.

Table 6.3: Parameters for the MBVD equivalent circuit of Figure 6-5.

R_m	$389\,\Omega$	R_{s1}	$2.89\,\Omega$
L_m	2.556×10^{-4} H	R_{s2}	$2.89\,\Omega$
C_m	9.539×10^{-17} F	C_{s1}	30 aF
C_f	47.39 fF	C_{s2}	30 aF
R_f	$8.4\,\Omega$	$f_\circ$	1.019 GHz
k^2_{eff}	0.2%	Q	4200

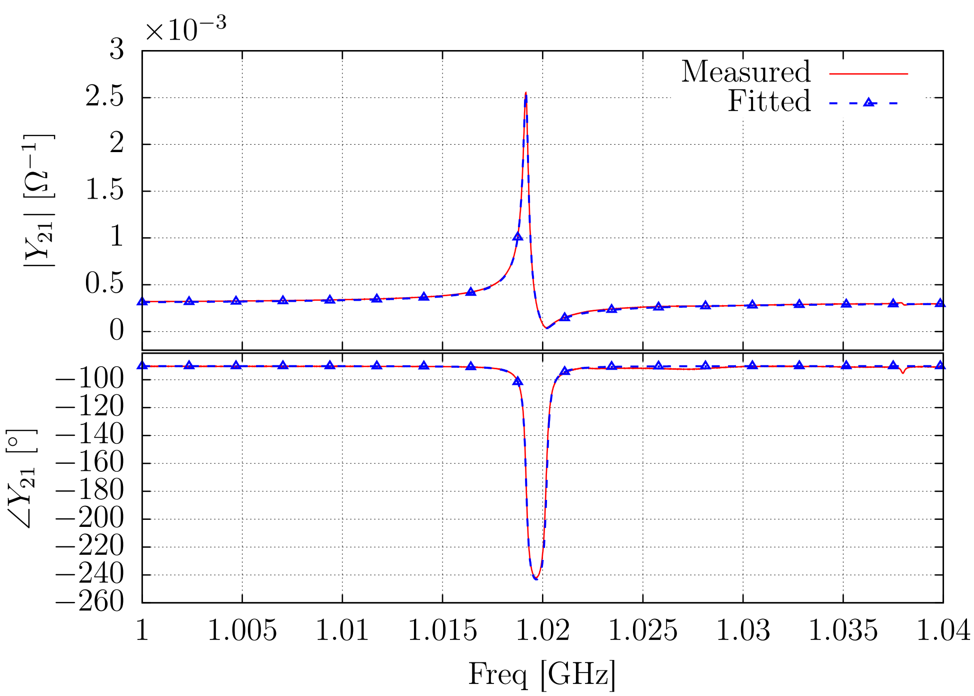


Figure 6-6: Lamb-mode resonator admittance Y_{21}, showing the measured response and the fitted results for the MBVD equivalent circuit.

6.3 High Electron Mobility Transistors (HEMTs)

The ability to build high-Q Lamb-mode resonators side by side with HEMTs and passives in GaN MMIC technology is the key advantage of L.C. Popa's process [44]. Monolithically integrated low-power low-phase noise oscillators and complete RF front-ends in GaN MMIC technologies become feasible. Monolithic integration significantly reduces the system size, parasitics, and overall power consumption.

6.3.1 HEMTs Model Requirements

In order to proceed with oscillator and circuit design in this process, accurate HEMT models are required. However, this being an in-house process developed at MTL, standard model cards were not available. A suitable model for the HEMTs must still be investigated.

Designing an oscillator circuit imposes stringent requirements on the transistor models. The oscillator circuit starts up in quiescent DC conditions. A small signal at a specific frequency (originating from thermal noise filtered by the resonator response) starts to get amplified as it travels around the oscillator loop. By virtue of the positive feedback, the oscillation grows until it is limited by the circuit's non-linearities*. Thus, oscillator operation involves quiescent DC bias conditions as well as small and large signal operation. For this reason, oscillator design requires a transistor model with the following characteristics:

- Spice-compatible model;
- Accurate DC biasing predictions;
- Well-defined small-signal AC parameters for start-up loop-gain analysis;
- Self-consistent current-charge formulation for accurate parasitics estimation;
- Well-behaved large-signal transient behavior for steady-state oscillation analysis;

*This is indeed the case for the oscillators considered here since the resonators power handling significantly exceeds the circuits linear range.

- Dispersion models, particularly for g_m;
- Noise models;
- Compatibility with periodic-steady-state (PSS) harmonic balance analysis.

In addition to the above-mentioned requirements, a physics-based model with a small number of parameters is preferred to simplify characterization and parameter extraction.

6.3.2 Available HEMT Models

Multiple HEMT models are available, including but not limited to Angelov-GaN, Curtice, EEHEMT, COMON, HSP, Dyna FET, and MVS [122–133]. A detailed comparison of these models has been studied in the literature [122, 123, 129]; a quick summary is provided in this subsection.

The Angelov model is a popular empirical model with 90 parameters, focusing on g_m non-linearities and lacking self-consistent charge-current formulation. The Curtice model is a semi-empirical model with 51 parameters, and earlier versions were not geometry scalable. The EEHEMT model is an empirical model with ~ 70 parameters, developed by Agilent. It has separate formulation for different operation regions as well as separate AC and DC behavior, yet it includes self-heating and dispersion models.

The COMON and HSP models both rely on a surface potential approach. Analytical expressions are obtained by self-consistently solving Poisson's equation and Schrödinger's equation in the channel quantum well, followed by applying a given transport model for the carriers. The result is a self-consistent channel current and terminal charge formulation.

The Dyna FET model relies on the training of an artificial neural network (ANN) with large signal non-linear measurements to extract a time-domain model. Dynamic trapping and de-trapping are also included in the Dyna FET model.

The aforementioned models usually involve very large numbers of parameters which complicates parameter extraction. Some of them also lack self-consistency, geometrical scaling, dispersion models and self-heating. This dictates the search for a more suitable model.

6.3.3 MIT Virtual Source (MVS) Model

The MIT Virtual Source (MVS) model developed at MIT [128–130] is a physics based model with self-consistent current-charge equations. Based on the injection velocity at the virtual source, the model performs well with short channel devices where quasi-ballistic transport is dominant. The model involves about 35 parameters, which greatly simplifies parameter extraction and fitting to measured devices when compared to the above-mentioned models. It also includes self-heating and robust access region model. However, the MVS model version used for this work lacked g_m dispersion and output kinks, both of which were strongly observed in the HEMTs fabricated at MTL. The MVS model has been augmented to include both effects as discussed below.

Keysight's ICCAP* was used to extract the parameters for the MVS model for different channel lengths. The parameter extraction procedure follows the procedure outlined in [129,130]. The process starts by extracting the gate capacitance from large AlGaN/GaN capacitors. Mobility and contact resistance are extracted from standard transmission line (TLM) measurements. Next, the HEMT DC characteristics are used to extract the threshold voltage, the subthreshold swing, and the DIBL. Virtual source velocity is then extracted with some adjustments to mobility and contact resistance to better fit the measurements.

DC output and subthreshold characteristics for $1\,\mu$m channel length and $50\,\mu$m width HEMT are shown in Figure 6-7. The augmented MVS model results closely match the DC measurement results, except for subthreshold $V_{gs} < -4\,$V as the gate leackage current is not modeled. It is also important to notice that this voltage is well below the threshold voltage and the HEMT is practically off. The HEMT shows a negligible g_m under this condition and is not useful for oscillator design. For this reason, this operation regime is neglected in the model.

*Formerly Agilent technologies

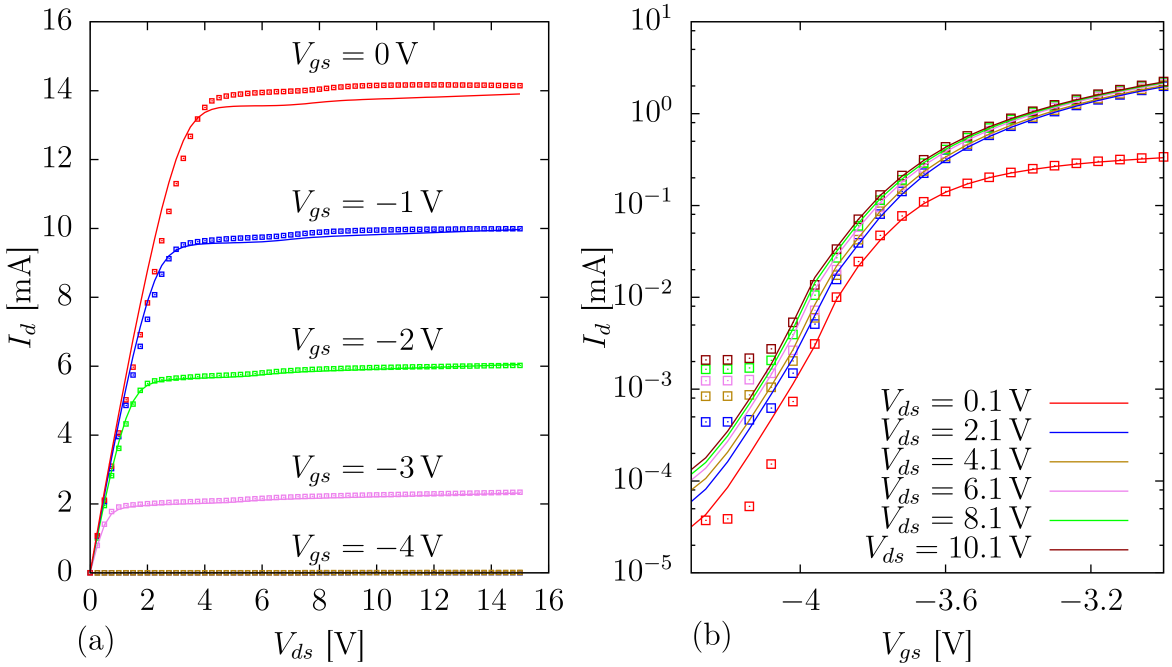


Figure 6-7: DC characteristics of a HEMT with 1 μm channel length and width of 50 μm: (a) output characteristics (b) transfer characteristics.

RF Current Collapse and g_m Dispersion

RF current collapse imposes a major limitation on HEMTs RF performance, manifested in gate-lag and drain-lag transients as well as g_m frequency dispersion [134–139]. This effect is mostly attributed to traps at the AlGaN surface, particularly in the ungated access region between the gate and drain [134, 135]. The energy band diagram for AlGaN/GaN heterostructure with an 2DEG formed is shown in Figure 6-8.a. The formation of 2DEG requires a positive charge on the AlGaN surface, which can be holes attracted to the AlGaN surface by negative polarization, ionized donor-like traps, or both. Charges accumulated on the AlGaN surface in the ungated access region between the gate and drain will modulate the 2DEG charge and the depletion region width, having the effect of a virtual gate that modulates drain current [135]. With electrons captured in the surface traps, the 2DEG density is reduced and might even completely disappear as in Figure 6-8.b. Traps occupancy is a function of bias conditions, electrons leaking from the gate, and hot carriers in the channel overcoming the AlGaN/GaN barrier. Current collapse and dispersion effects characterized by time constants on the order of $10^{-2} - 1\,\mathrm{s}$ can mostly be attributed to captured electrons in surface traps. This effect can be reduced by the proper passivation of AlGaN surface with SiN. Moreover, hole traps can capture holes from the AlGaN layers contributing to the reduction of the 2DEG charge. This mechanism is responsible for current collapse and dispersion with much shorter time constants on the order of 10 to $100\,\mu\mathrm{s}$ [134].

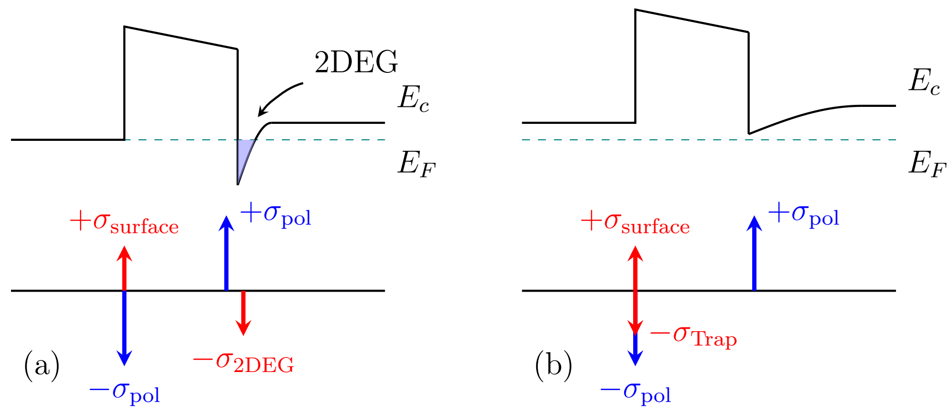


Figure 6-8: Energy band diagram highlighting the conduction band edge and the Fermi level for AlGaN/GaN heterostructure. It clearly shows (a) 2DEG with high concentration, compensated by positive charge on the AlGaN surface and (b) 2DEG disappearing with enlargement of the GaN depletion region as a result negative surface resulting from captured electron.

For oscillator design, HEMTs will be operating in a continuous wave mode rather than a switching regime. Thus, g_m dispersion becomes a primary concern to ensure that the oscillator has enough startup gain. It is important to incorporate such effects in the MVS model to be used for oscillator design. However, it is important to note that such oscillators operate at a single frequency, with oscillation startup being mostly small signal operation. Moreover, for GHz oscillators, current collapse mechanisms are at least 4 orders of magnitude slower than the oscillation period. Hence, for the purposes of our application, it is enough to consider g_m dispersion between DC and GHz frequencies without focusing on the exact time constant of current collapse.

g_m dispersion is incorporated into the MVS model by first isolating the RF component of the drain current with a high pass R-C filter. The RF component of the drain current i_{D-RF} is given by

$$i_{D-RF} = i_{D\circ} \frac{sRC}{1 + sRC}, \tag{6.2}$$

where $i_{D\circ}$ is the total drain current without dispersion. For the targeted oscillator design, the R-C time constant is orders of magnitude slower than the frequency of interest. Next, based on the observation of g_m measurement, a gate voltage bias dependence is selected in the form

$$x_{gm} = x_{\circ-gm} \exp\left(-\left(\frac{V_{gs} - V_{gm}}{\gamma_{gm}}\right)^2\right), \tag{6.3}$$

where $x_{\circ-gm}$, V_{gm} and γ_{gm} are fitting parameters. The total instantaneous HEMT drain current is calculated as

$$i_D = i_{D\circ} - x_{gm} \times i_{D-RF}. \tag{6.4}$$

This simple g_m dispersion formulation was implemented in Verilog-A and is fully compatible with PSS analysis, as it doesn't involve any hidden states or discontinuous definitions.

Model-fitted g_m is compared to measurement results for both DC and RF operation in Figure 6-9 and Figure 6-10, respectively. g_m dispersion behavior is evident from the measurements, and the model does a good job capturing such behavior. It is important to note that RF g_m was measured with -36 dBm of power to avoid disturbing the DC operating point of the transistor near subthreshold.

Output Kinks

As seen in Figure 6-7 on page 192, kinks are observed in the output characteristics of our HEMTs. The kink phenomenon is fairly complex and is attributable either to intraband impact ionization of acceptor-like states or electron trapping in donor-like states [140–143].

The direct effect of such kinks on circuit design is the sudden increase in the output

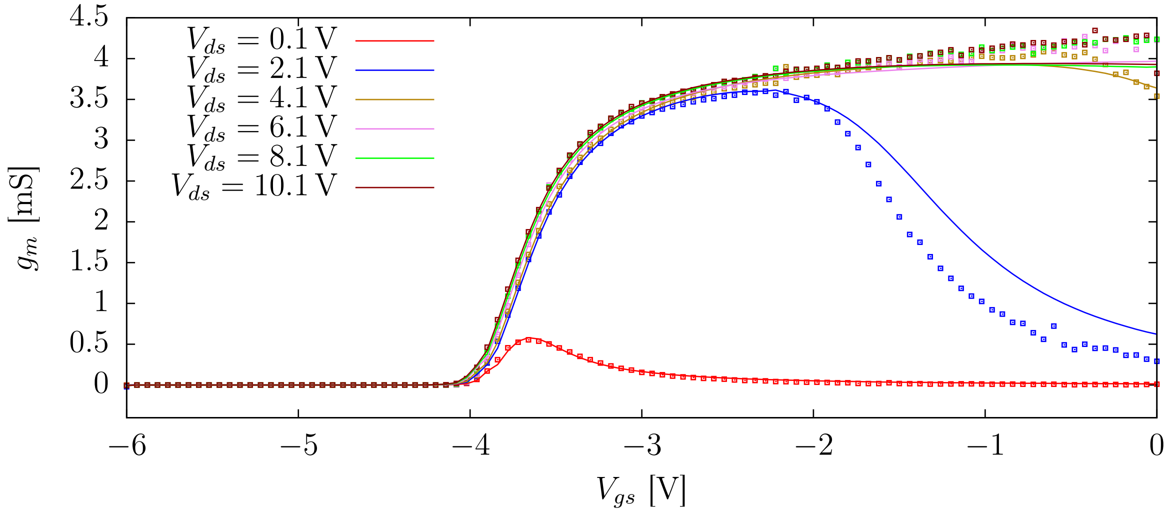


Figure 6-9: DC g_m for a HEMT with $1\,\mu$m gate length and $50\,\mu$m width showing measured characteristics (squares) and fitted model results (solid lines).

conductance of the transistor as seen in Figure 6-11 on page 197. While such kinks might have slow time constants (as they are related to deep traps), including their output conductance effect in the model at all frequencies will be a very conservative strategy.

The kink is incorporated in the MVS model in the DIBL factor δ. First, the dependence of the location of the kink on V_{gs} is captured by

$$V_{kf} = \ln\left[1 + \exp\left(\frac{V_{gs} - V_t}{V_{sk}}\right)\right], \tag{6.5}$$

where V_{sk} is a fitting parameter. The amplitude of the kink A_k is captured by

$$A_k = \frac{1}{2}\left(1 + \tanh\left[\eta_k \left(V_{gs} - V_t\right)\right]\right), \tag{6.6}$$

where η_k is a fitting parameter. Finally, the kink effect is added to the DIBL factor of the MVS model as

$$\psi_k = 1 + \tanh\left[\beta_k \left(V_{ds} - V_{kf} - V_k\right)\right], \tag{6.7a}$$

$$\delta = \left(\delta_1 + \delta_k \times A_k \times \psi_k - V_{\text{sat-DIBL}} \times \delta_2\right) \times V_{ds}, \tag{6.7b}$$

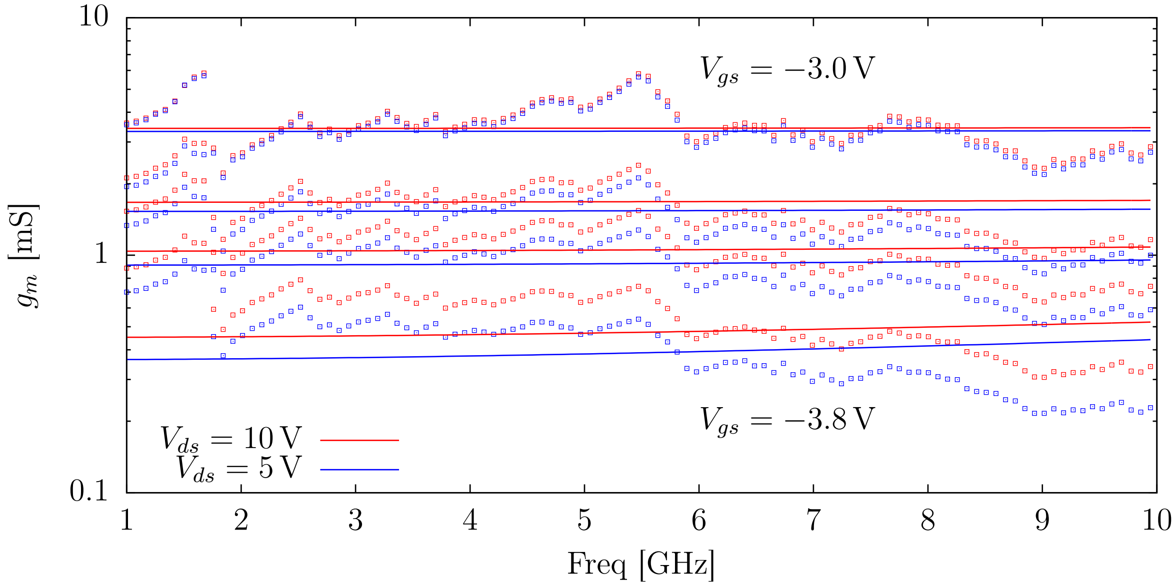


Figure 6-10: RF g_m for a HEMT with $1\,\mu$m gate length and $50\,\mu$m width, showing measured (squares) vs fitted model (solid lines) results for V_{gs} ={-3.8 V, -3.7 V, -3.6 V, -3.0 V} and V_{ds} ={10 V, 5 V}.

where β_k, V_k and δ_k are fitting parameters; δ_1, δ_2 and $V_{\text{sat-DIBL}}$ are MVS model parameters.

The different kink model parameters are extracted using ICCAP for HEMTs of different dimensions. Figure 6-11 on the next page compares measurements results and extracted model for a HEMT with $1\,\mu$m gate length and $50\,\mu$m width. The model reasonably captures the kink effect. As mentioned above, such increase in output conductance is set to be frequency-independent assuming a conservative design strategy. Time constant can be added to the model in a similar fashion to that described for the g_m-dispersion.

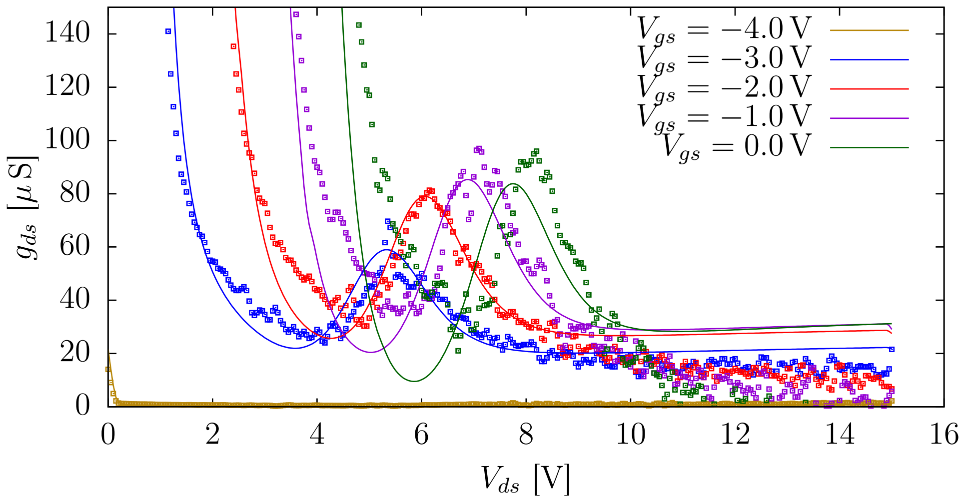


Figure 6-11: Output conductance for a HEMT with 1 μm gate length and 50 μm width, showing measured (squares) versus fitted model (solid lines).

6.4 Colpitts and Pierce Oscillators

The unique ability to fabricate high-Q MEMS resonators side-by-side with HEMTs, enables low power and low phase noise monolithic oscillators in GaN MMIC technologies. In this section, monolithic Colpitts and Pierce oscillators are demonstrated.

6.4.1 HEMTs DC Biasing

One of the challenges in implementing a MEMS-based oscillator with HEMTs is their strict DC biasing requirements. Unlike standard CMOS MOSFETs, HEMTs are depletion mode devices that are normally on. Normal operation will require the gate Schottky diode to be reverse biased, imposing a negative V_{gs}. Moreover, for low-power oscillator design, HEMTs must be operated in sub-threshold regime for a high g_m/I_d, V_{gs} that is even more negative.

Unlike p-MOS devices that require both a negative V_{gs} and a negative V_{ds}, HEMTs require a negative V_{gs} and a positive V_{ds} to operate. These requirements complicate the HEMTs DC biasing scheme. They also makes simple circuit building blocks like diode-connected MOS transistors infeasible. This directly impacts our ability to construct a current mirror with HEMTs while maintaining a simple circuit architecture.

Owing to the simplicity of both the Pierce and Colpitts oscillator circuits, they can still be realized with HEMTs while operating from a single supply rail. The key enabler of such circuits is the use of a source resistance R_s while biasing the gate to ground through a resistor R_g as shown in Figure 6-12. This creates the negative V_{gs} required while maintaining positive V_{ds}. However, it is important to note that this naïve biasing scheme will be susceptible to process variations. The lack of simple current mirrors complicates the realization of constant-g_m biasing [144] that is rather preferable in such oscillators designs to desensitize the oscillator loop-gain against process variations.

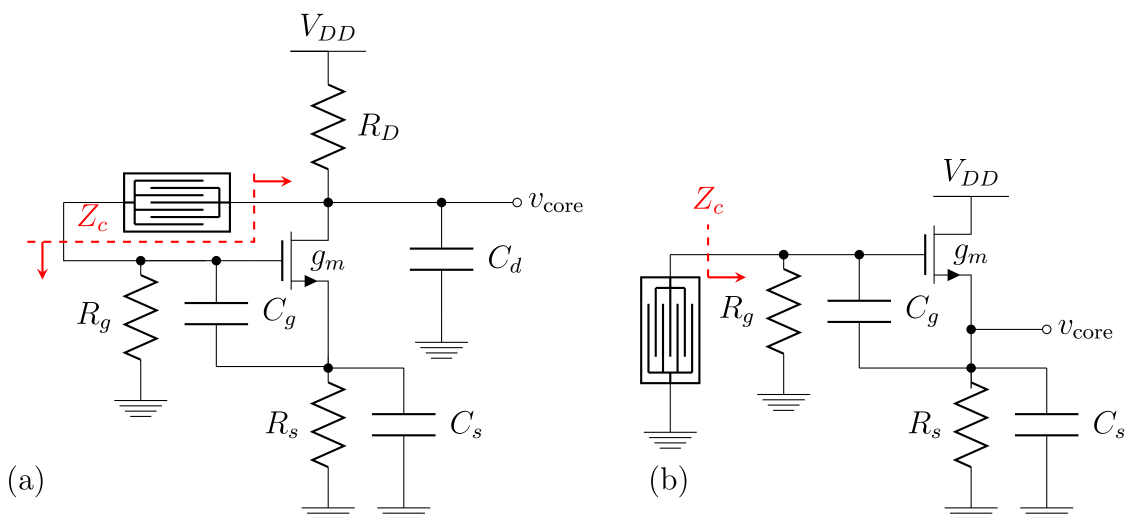


Figure 6-12: Typical circuit schematic for (a) Pierce and (b) Colpitts oscillators, showing the source resistance used to provide DC biasing for the HEMT.

Figure 6-12 suggests that the Colpitts oscillator can operate from a lower supply voltage than the Pierce oscillator due to the lack of drain resistance R_D. Moreover, the Colpitts oscillator presented here provides zero DC bias for the resonator. This is beneficial to avoid pre-straining the oscillator with the DC bias and also simplifies the circuit in the case of using 2DEG electrodes for resonator transduction, for example*.

*Extended discussion about resonators using 2DEG electrodes can be found in [44]

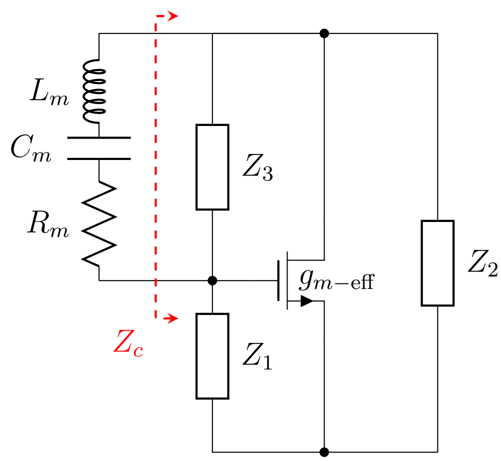


Figure 6-13: Generic 3-point oscillator small-signal equivalent circuit.

Table 6.4: 3-point oscillator equivalent circuit parameters for Pierce and Colpitts

Equivalent Circuit	Pierce Oscillator	Colpitts Oscillator
$g_{m-\text{eff}}$	$\frac{g_m}{1+g_m\,(R_s \| 1/s\,C_s)}$	g_m
Z_1	$R_g \parallel 1/s\,C_g$	$1/s\,(C_g + C_{gs})$
Z_2	$R_D \parallel 1/s\,C_{ds}$	$R_s \parallel 1/s\,(C_s + C_{ds})$
Z_3	$(R_f + 1/sC_f) \parallel 1/sC_{gd}$	$(R_f + 1/sC_f) \parallel R_g$

6.4.2 Three-Point Oscillator

Critical Transconductance and Power Consumption

Both Pierce and Colpitts oscillators can be analyzed and designed according to classical three-point oscillator theory [145]. Both circuits can be reorganized into the generic 3-point oscillator equivalent circuit of Figure 6-13 with equivalent circuit parameters as shown in Table 6.4. The resonator parasitics are lumped into the equivalent impedance Z_c. For the Pierce oscillator, the degeneration resistance R_s results in a decreased effective $g_{m-\text{eff}}$. It is hence desirable to have the time constant R_sC_s be orders of magnitudes larger than the oscillator period $(2\pi/\omega_\circ)$. Under this assumption $g_{m-\text{eff}}$ approaches g_m for the Pierce oscillator.

The motional impedance Z_m is given by

$$Z_m = R_m + \frac{1}{j\omega_\circ C_m} + j\omega_\circ L_m = R_m + \frac{j}{\omega_m C_m}\left(\frac{\omega_\circ}{\omega_m} - \frac{\omega_m}{\omega_\circ}\right), \tag{6.8}$$

where $\omega_m = 1/\sqrt{L_m C_m}$ is the series resonance frequency of the motional branch and

$\omega_\circ$ is the actual oscillation frequency. Define the frequency pulling p as

$$p = \frac{\omega_\circ - \omega_m}{\omega_m}. \tag{6.9}$$

For most oscillator designs, the pulling is very small (well below 1%), allowing the motional impedance to be approximated as

$$Z_m = R_m + \frac{j}{\omega_m C_m}\left(1 + p - \frac{1}{1+p}\right) \approx R_m + j\frac{2p}{\omega_m C_m}. \tag{6.10}$$

Assuming small signal operation, the condition for sustained oscillation is*

$$Z_m + Z_c = 0, \tag{6.11}$$

or

$$R_m + \mathrm{Re}\{Z_c\} = 0 \qquad \text{and} \qquad \frac{2p}{\omega_m C_m} + \mathrm{Im}\{Z_c\} = 0. \tag{6.12}$$

This condition suggests that it is necessary to have a negative $\mathrm{Re}\{Z_c\}$ equal to $-R_m$ (which actually controls the oscillator loop gain) for sustained oscillation. The transistor, along with the impedances Z_1 to Z_3, creates such negative resistance. It also indicates that the imaginary part $\mathrm{Im}\{Z_c\}$ is responsible for frequency pulling. Startup condition, startup time, and frequency pulling can be analyzed by studying the equivalent circuit impedance Z_c.

The core impedance Z_c can be obtained from the circuit of Figure 6-13 on the preceding page as

$$Z_c = \frac{Z_1 Z_3 + Z_2 Z_3 + g_{m-\mathrm{eff}} Z_1 Z_2 Z_3}{Z_1 + Z_2 + Z_3 + g_{m-\mathrm{eff}} Z_1 Z_2}. \tag{6.13}$$

Figure 6-14 on the next page shows the locus of Z_c in the complex plane as a function of g_m, which takes the form of a circle. As the transistor g_m is increased from 0 all the way to ∞, Z_c moves along the left half of this circle [145]. The center and radius of the circle are set by the values of Z_1, Z_2 and Z_3. To satisfy the conditions

*This is the same as Barkhausen criteria by considering the positive feedback loop in this circuit, where a voltage is imposed on Z_m by Z_c, which generates a current that flows through Z_c and imposes the voltage on Z_m.

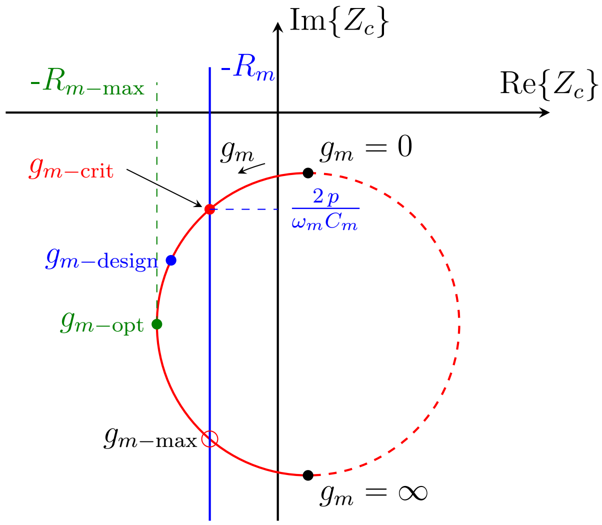


Figure 6-14: Locus of Z_c as function of g_m in the complex plane showing critical, optimal and maximum transconductance values for the oscillator core transistor.

(6.12), the left half of the circle is required to intercept the line $\mathrm{Re}\{Z_c\} = -R_m$. This interception marks $g_{m-\mathrm{crit}}$, the smallest g_m required for a sustained oscillation. In fact, it is required to have $g_m > g_{m-\mathrm{crit}}$ to guarantee sufficient loop gain for oscillator startup.

Assuming R_g and R_D are very large, whereas R_s and R_f are negligible, the 3-point oscillator circuit of Figure 6-13 on page 199 reduces to a lossless capacitive core circuit. Although these assumptions might not hold, the result of the simplification gives great insight into the oscillator operation. The maximum negative resistance available from the core circuit is found to be [145]

$$\mathrm{Re}\{Z_c\}_{\max} = -\frac{1}{2\omega_\circ C_3(1 + C_3\frac{C_1+C_2}{C_1C_2})}, \tag{6.14}$$

where C_1, C_2 and C_3 are the capacitances corresponding to Z_1, Z_2 and Z_3, respectively. This maximum negative resistance is achieved with $g_m = g_{m-\mathrm{opt}}$

$$g_m = g_{m-\mathrm{opt}} = \omega_\circ \left(\frac{C_1C_2 + C_2C_3 + C_1C_3}{C_3}\right). \tag{6.15}$$

Hence, the condition for a feasible oscillator design becomes

$$\mathrm{Re}\{Z_c\}_{\max} < -R_m \Rightarrow \frac{QC_m}{C_3} > 2\left(1 + C_3\frac{C_1 + C_2}{C_1C_2}\right). \tag{6.16}$$

The electromechanical coupling coefficient k_{eff}^2 can be approximated as [146, 147]

$$k_{eff}^2 \approx \frac{\pi^2}{8}\frac{C_m}{C_f}, \tag{6.17}$$

when its value is sufficiently small (typically on the order of 1%). By noting that $C_3 \approx C_f$, the condition (6.16) for feasible oscillator becomes

$$k_{eff}^2 Q > \frac{\pi^2}{4}\left(1 + C_3\frac{C_1 + C_2}{C_1C_2}\right), \tag{6.18}$$

where $k_{eff}^2 Q$ is recognized as a resonator figure of merit. The critical transconductance $g_{m-\mathrm{crit}}$ required can be approximated as [145]

$$g_{m-\mathrm{crit}} \approx \frac{\omega_\circ}{QC_m}\frac{(C_1C_2 + C_2C_3 + C_3C_1)^2}{C_1C_2}. \tag{6.19}$$

Still assuming that $C_3 \approx C_f$, $g_{m-\mathrm{crit}}$ can be further expressed as

$$g_{m-\mathrm{crit}} \approx \frac{\pi^2}{8}\frac{\omega_\circ}{k_{eff}^2 Q}\frac{(C_1C_2 + C_2C_3 + C_3C_1)^2}{C_1C_2C_3}. \tag{6.20}$$

This relation directly correlates the minimum transconductance required $g_{m-\mathrm{crit}}$ to the resonator figure of merit $k_{eff}^2 Q$, the operating frequency $\omega_\circ$, and the electrical capacitances C_1, C_2 and C_3. With the HEMT selected to operate with a given $g_m/I_d = (g_m/I_d)_\circ$, an estimation of the oscillator minimum current consumption $I_{\text{DC-min}}$ is directly available from (6.20) as

$$I_{\text{DC-min}} \approx \frac{\pi^2/8}{(g_m/I_d)_\circ}\frac{\omega_\circ}{k_{eff}^2 Q}\frac{(C_1C_2 + C_2C_3 + C_3C_1)^2}{C_1C_2C_3}. \tag{6.21}$$

which can also be expressed in terms of the resonator $f_\circ \cdot Q$ as

$$I_{\text{DC-min}} \approx \frac{\pi/16}{(g_m/I_d)_\circ} \frac{\omega_\circ^2}{k_{eff}^2 \times f_\circ \times Q} \frac{(C_1C_2 + C_2C_3 + C_3C_1)^2}{C_1C_2C_3}, \tag{6.22}$$

where small pulling was assumed, resulting in $\omega_\circ \approx \omega_m$. At low-frequencies with acoustic wavelength considerably larger than the phonons' mean free path, the quality factor is limited by phonon relaxation in the Akheiser regime, resulting in a constant $f_\circ \cdot Q$ product for the resonator [82]. In this regime, the oscillator power consumption scales up quadratically with the operating frequency. However, at high frequencies with much shorter wavelengths, phonon relaxation is limited by the Landau-Rumer regime, resulting in an $f_\circ \cdot Q$ product that scales linearly with frequency [82]. In the Landau-Rumer regime, the oscillator power consumption scales linearly with the operating frequency. In both regimes the power consumption scales linearly with the electrical capacitance* and is inversely proportional to k_{eff}^2. This discussion suggests that for low power oscillator design, $k_{eff}^2 \times f_\circ \times Q$ should be used as a resonator figure of merit in the Akheiser regime, whereas $k_{eff}^2 \times Q$ is a more suitable figure of merit in the Landau-Rumer regime.

Frequency Pulling

The oscillator frequency pulling p depends on the resonator motional resistance as well as the circuit capacitances (the latter controls the circle position and radius). To minimize such dependence and maximize frequency stability, it is desirable to satisfy (6.16) with a large margin as well. Under this condition, p becomes independent of R_m and is approximately fixed at its value with $g_m = 0$. The frequency pulling is found to be [145]

$$p = \frac{C_m}{2(C_3 + \frac{C_1C_2}{C_1+C_2})}. \tag{6.23}$$

*Assuming all capacitors are equally scaled; otherwise, the dependence is sublinear.

This can also be combined with (6.19) to yield

$$\begin{aligned} g_{m-\text{crit}} \times p^2 &= \frac{\omega_\circ C_m}{Q} \frac{(C_1+C_2)^2}{4C_1C_2} \\ I_{\text{DC-min}} \times p^2 &= \omega_\circ \frac{8/\pi^2}{(g_m/I_d)_\circ} \frac{k_{eff}^2}{Q} \frac{C_3(C_1+C_2)^2}{4C_1C_2}. \end{aligned} \tag{6.24}$$

This highlights a trade-off between power consumption and frequency stability [145]. Also, combining (6.20) and (6.24) yields

$$p > \frac{\pi^2}{8} \frac{C_m}{k_{eff}^2 Q} \frac{C_1+C_2}{C_1C_2} \qquad \text{or} \qquad p > \frac{1}{Q} C_f \frac{C_1+C_2}{C_1C_2}, \tag{6.25}$$

which sets a lower bound for the frequency pulling p relating to the resonator quality factor Q and feed-through capacitance C_f.

Startup Time

The minimum startup time τ_{min} is achieved for the transconductance $g_{m-\text{opt}}$ and is given by [145]

$$\tau_{min} = \frac{2C_f}{\omega_\circ C_m} \left(1 + \frac{C_f}{C_1} + \frac{C_f}{C_2}\right), \tag{6.26}$$

which can be rearranged to yield

$$\tau_{min} > \frac{2C_f}{\omega_\circ C_m} \Rightarrow \tau_{min} > \frac{\pi^2}{4} \frac{1}{k_{eff}^2} \frac{1}{\omega_\circ}. \tag{6.27}$$

This clearly relates the oscillator startup time to k_{eff}^2. With typical k_{eff}^2 of 0.2% for our resonator, these oscillators' startup time is on the order of 1,000 cycles.

6.4.3 Oscillator Design

As outlined in the previous subsection, the first step in a Pierce oscillator and Colpitts oscillator design is to find the critical g_m required for oscillation startup, as well as $g_{m-\text{opt}}$ and estimated startup time. This is readily calculated from (6.13). Using the full forms for $g_{m-\text{eff}}$, Z_1, Z_2 and Z_3 and taking into consideration all resonator parasitics, accurate $g_{m-\text{crit}}$ calculation is feasible. The resulting transcendal equations do not admit analytical solution and have to be solved numerically.

A Mathematica [148] graphical user interface (GUI), shown in Figure 6-15, has been designed to carry out these calculations. Fitted equivalent circuit parameters for all measured resonators are made available for direct access from the GUI, whereas a local database is used to store different oscillator designs. The GUI also provides visualization for the complex impedance plane of Figure 6-14 on page 201.

As an in-house process, large variations were assumed, and hence multiple versions of Pierce and Colpitts oscillators were implemented with different degrees of over-design. Three main oscillators are considered here*:

- **Pierce A:** designed for a maximum motional resistance of 600 Ω with 1 mW power consumption;
- **Colpitts:** designed for a maximum motional resistance of 1 kΩ with 1.5 mW power consumption;
- **Pierce B:** designed for a maximum motional resistance of 2 kΩ with 4 mW power consumption.

The design process starts by selecting a resonator and reasonable values for the capacitors C_1 and C_2 that are small enough to reduce the power consumption but sufficiently large to reduce the effect of the HEMTs parasitic capacitances. A critical transconductance is found from the GUI of Figure 6-15. A design transconductance that is twice this value is selected. HEMTs are operated in the subthreshold regime to maximize their g_m/I_d efficiency. Based on the required design g_m and g_m/I_d, the HEMT current and dimensions are selected [149]. HEMTs parasitics are then back annotated in the design GUI to assess their performance on the design, iterating as necessary. The design parameters for the three oscillators are listed in Table 6.5 on page 209. Pierce B in particular is not operated in subthreshold to reduce para-

*These oscillators were in fact designed and fabricated before incorporating the g_m-dispersion and the output kinks into the HEMTs model.

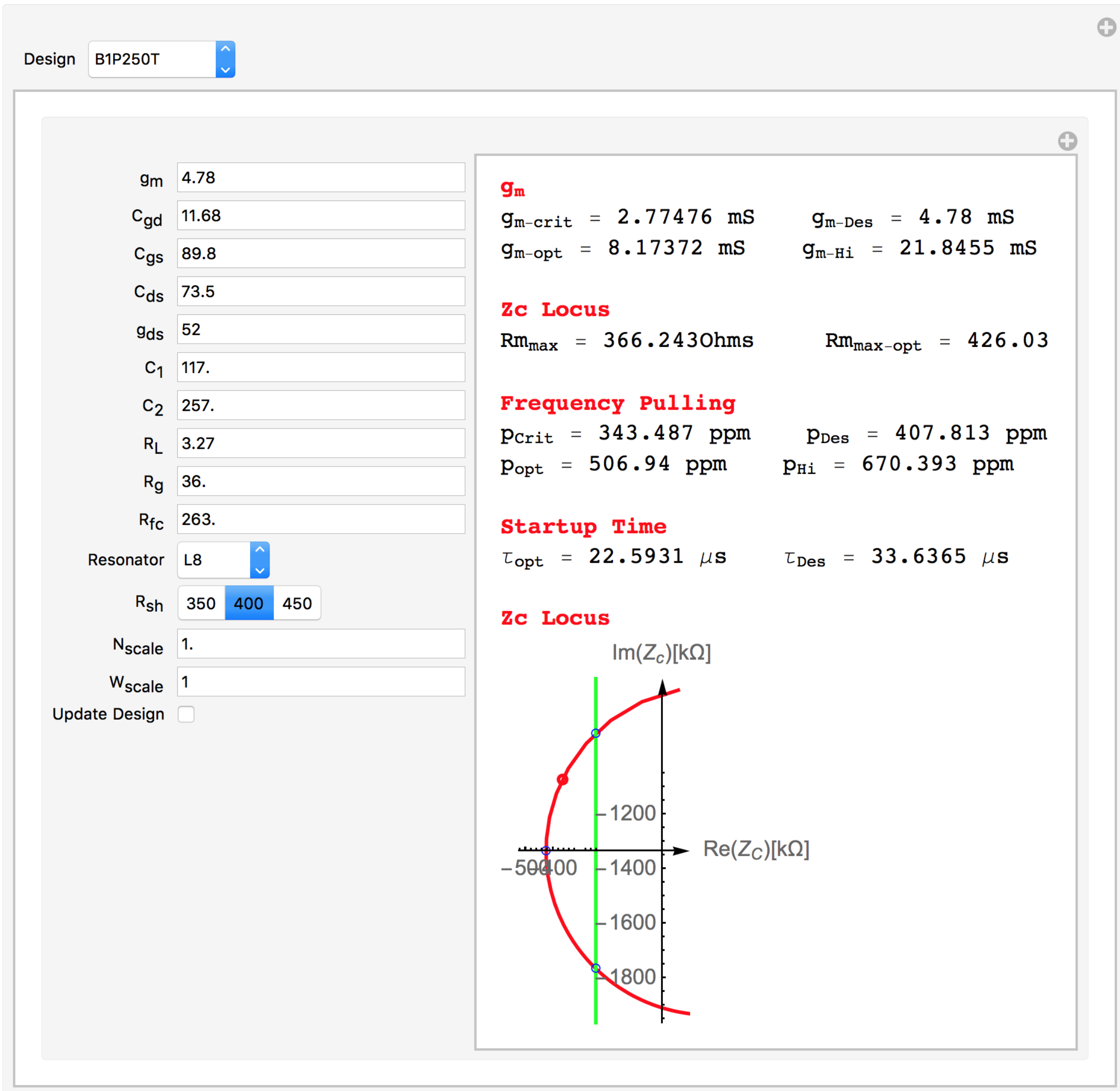


Figure 6-15: Mathematica GUI to help design Pierce and Colpitts oscillators. The GUI has access to all resonator models available and includes a local database to save all the different oscillator designs.

sitic capacitance. The capacitance C_g is even omitted, relying on the HEMT C_{gs} capacitance to provide C_1 for the circuit.

The three oscillator variations have been fabricated and tested at MTL, and the results are reported in §6.5. It is important to highlight the fact that the oscillators were fabricated on a GaN on $< 111 >$-Si wafer with different stress levels than that used to characterize the HEMT models during initial design phase. The HEMTs in the fabricated oscillators suffered a severe degradation of their DC and RF performance. Table 6.6 on page 209 compares the oscillators' performance parameters estimated from original models as opposed to models fitted to HEMTs and passives fabricated on the same die as the oscillators. The simulation results reported hereafter are based on the newly fitted HEMTs and passives*.

Simulated loop gain for all oscillators is shown in Figure 6-16 on the following page. The degradation in HEMT performance significantly degraded the loop-gain, but not so much as to prevent oscillator startup. The startup behavior of all oscillators is shown in figures 6-17, 6-18 and 6-19. The steady state oscillation period is shown in Figure 6-20 on page 211. Phase noise simulation results are shown in Figure 6-21 on page 212. Simulation clearly shows the over-designed Pierce B as having larger loop gain, start-up time that is 69% shorter than the other oscillators and 10 dB better phase noise performance.

*Adjustments had to be made for individual oscillators in accordance with across-chip fabrication variations to obtain correct fitting for the RF output power of the oscillators.

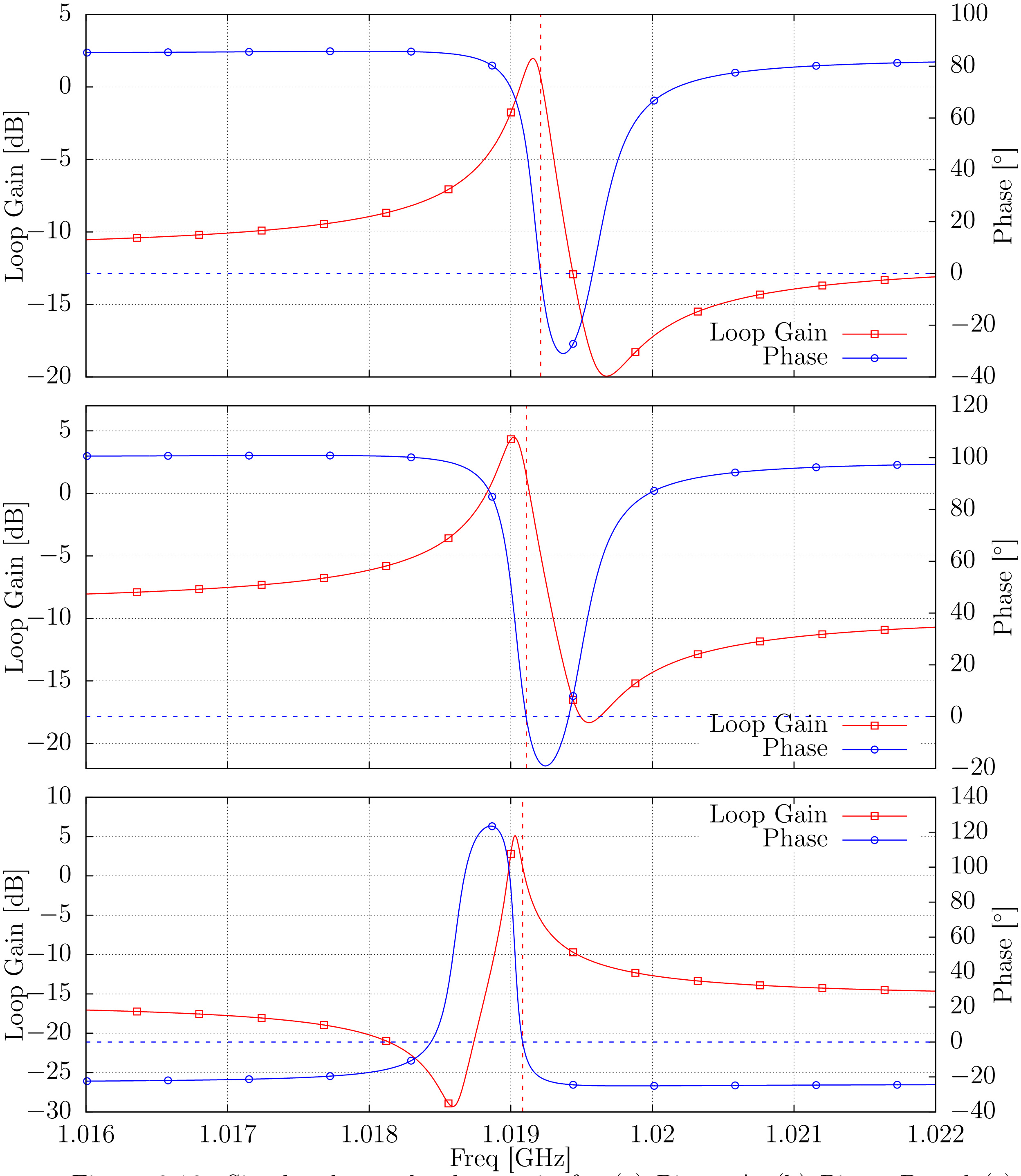


Figure 6-16: Simulated complex loop gain for (a) Pierce A, (b) Pierce B and (c) Colpitts oscillators implemented in this work.

Table 6.5: Design parameters for the oscillators considered.

	Pierce A	Pierce B	Colpitts
HEMT w/l	50/1	60/1	50/1
R_D (kΩ)	30	10	-
R_s (kΩ)	27.9	5.7	14.7
R_g (kΩ)	54	54	69.5
C_s (fF)	241	289	117
C_g (fF)	45	-	26.2
C_d (fF)	52	105	-

Table 6.6: Comparison between oscillator performance parameters as designed and as fitted to measurements due to the different stress-engineering in the GaN on silicon wafer.

	As designed			Fitted to measurements		
	Pierce A	Pierce B	Colpitts	Pierce A	Pierce B	Colpitts
g_m(mS)	1.354	3.91	1.898	0.756	2.132	1.343
Loop Gain (dB)	13	6.2	9.2	0.7	1.5	1.18
I_{DC} (μA)	135	630	266	93	390	210
$V_{\text{core-pp}}$ (V)	2.124	2.79	1.22	0.328	1.435	0.402
$V_{\text{out-pp}}$ (mV)	520	569	274	40	165	46

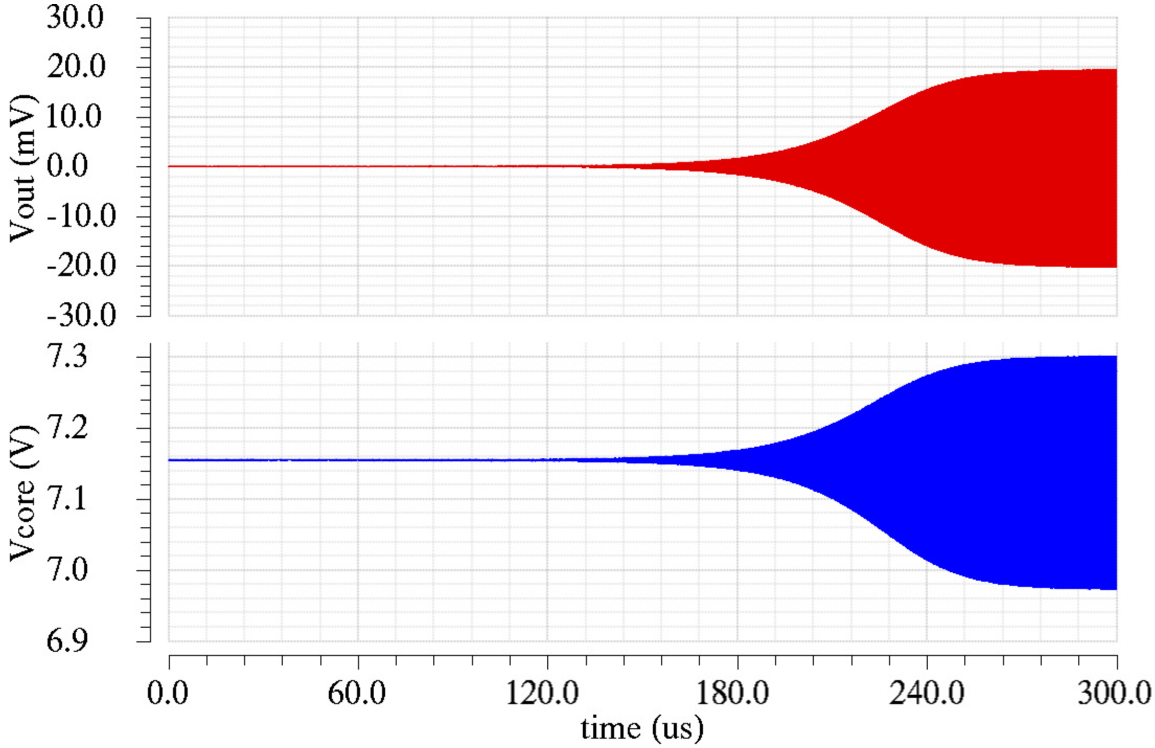


Figure 6-17: Startup behavior for oscillator Pierce A showing output after the buffer (V_{out}) and core output voltage (V_{core}).

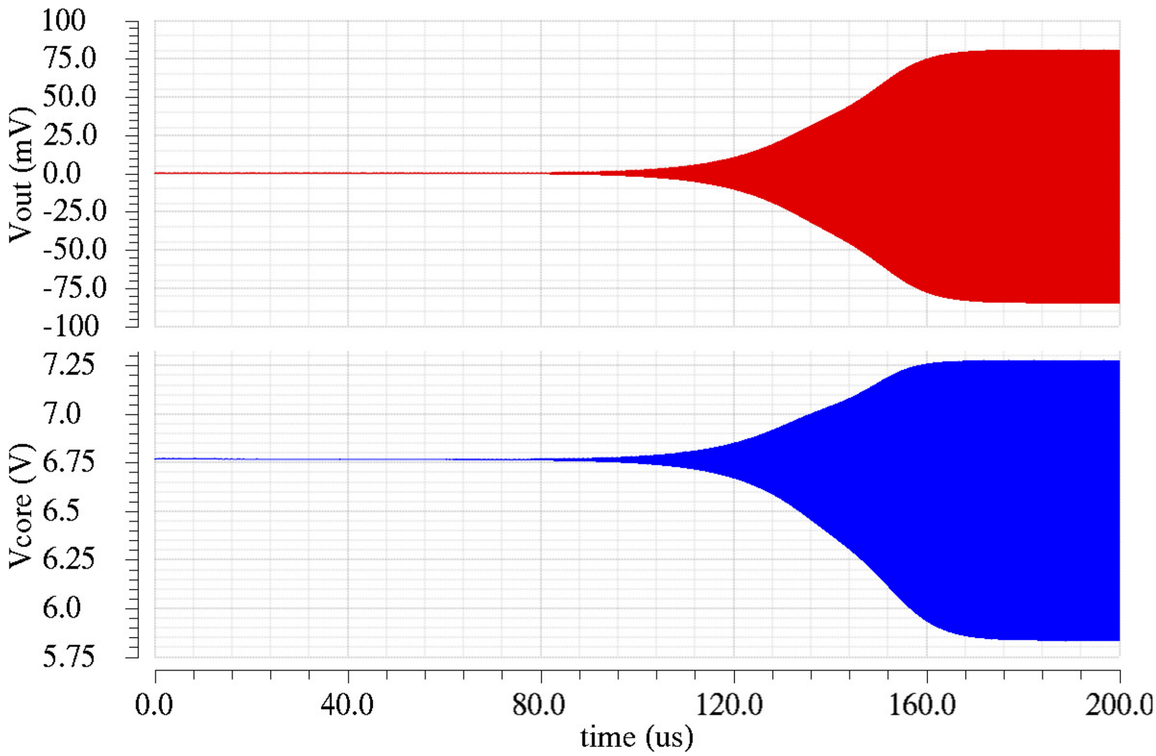


Figure 6-18: Startup behavior for oscillator Pierce B showing output after the buffer (V_{out}) and core output voltage (V_{core}).

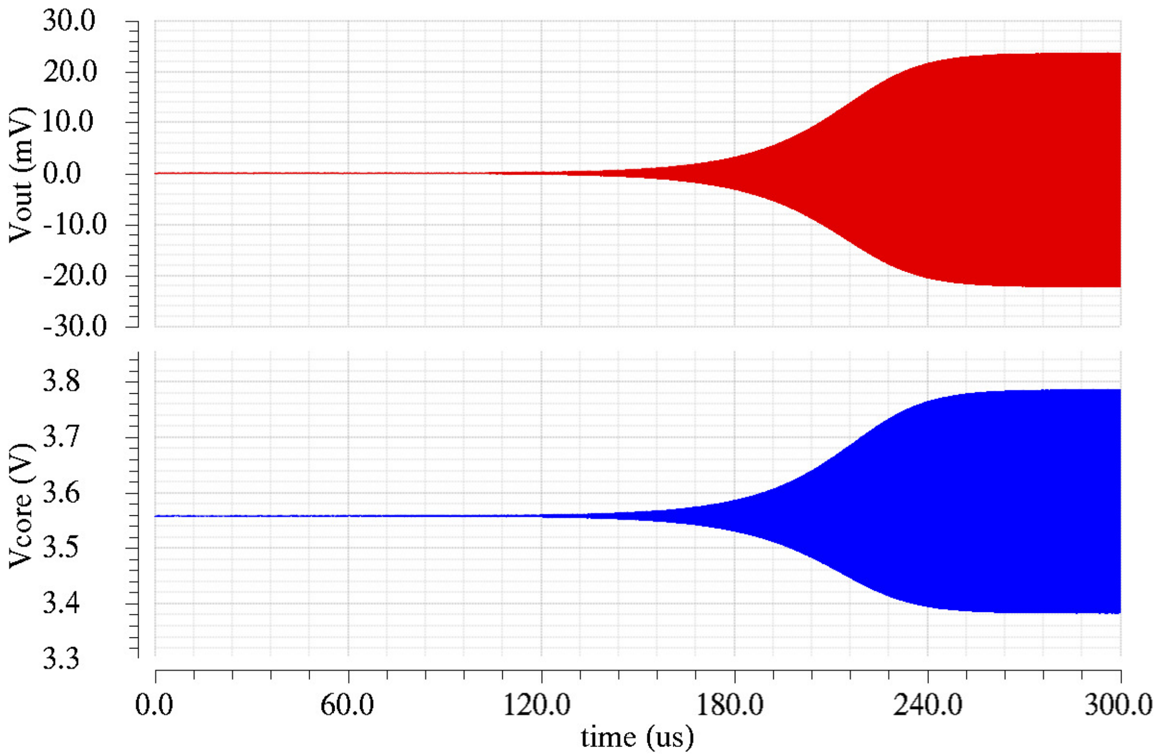


Figure 6-19: Startup behavior for Colpitts oscillator showing output after the buffer (V_{out}) and core output voltage (V_{core}).

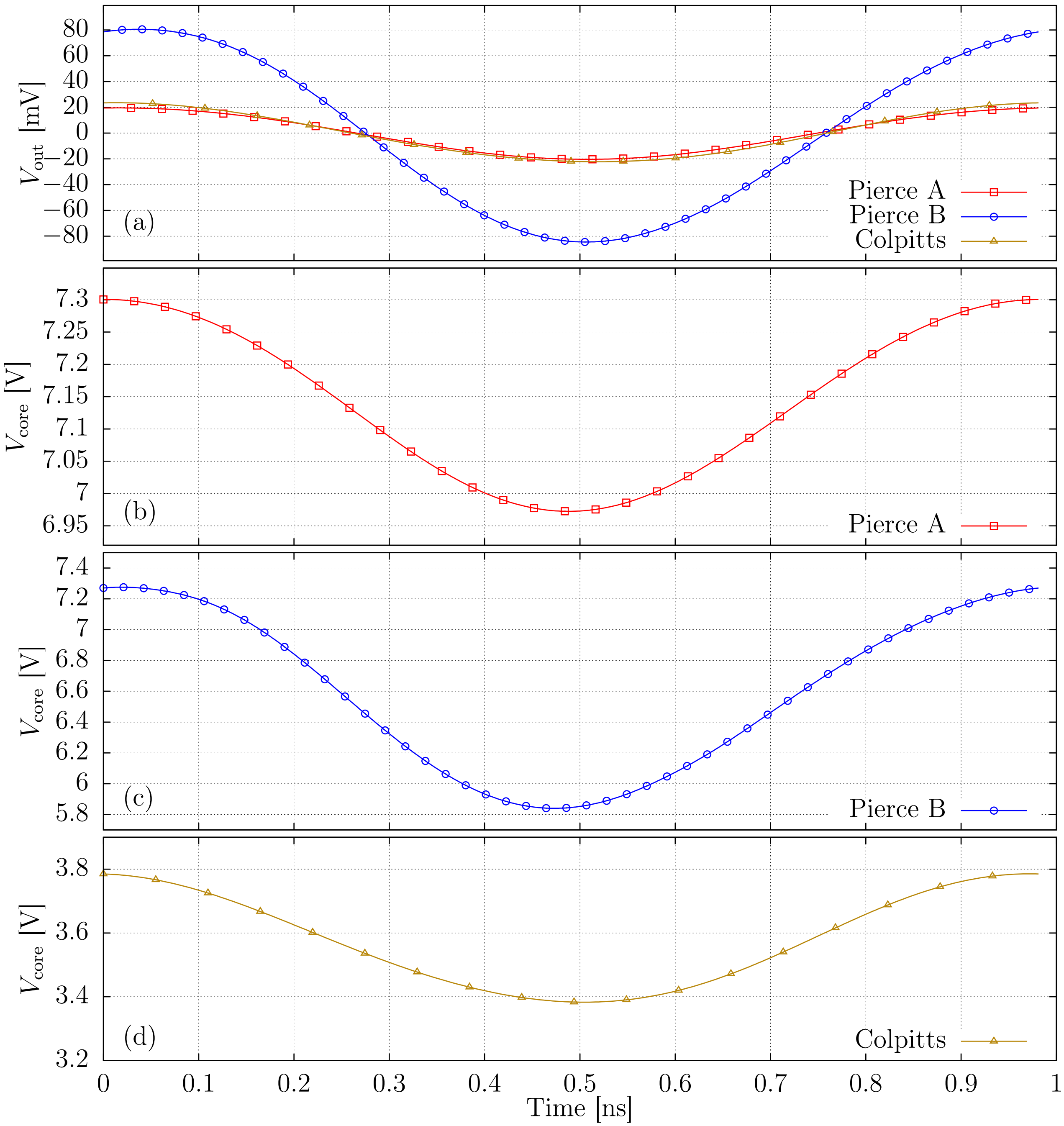


Figure 6-20: Oscillator steady state output waveforms, showing (a) the buffer output for all 3 oscillators considered, and the core output voltage for (b) Pierce A oscillator, (c) Pierce B oscillator and (d) Colpitts oscillator.

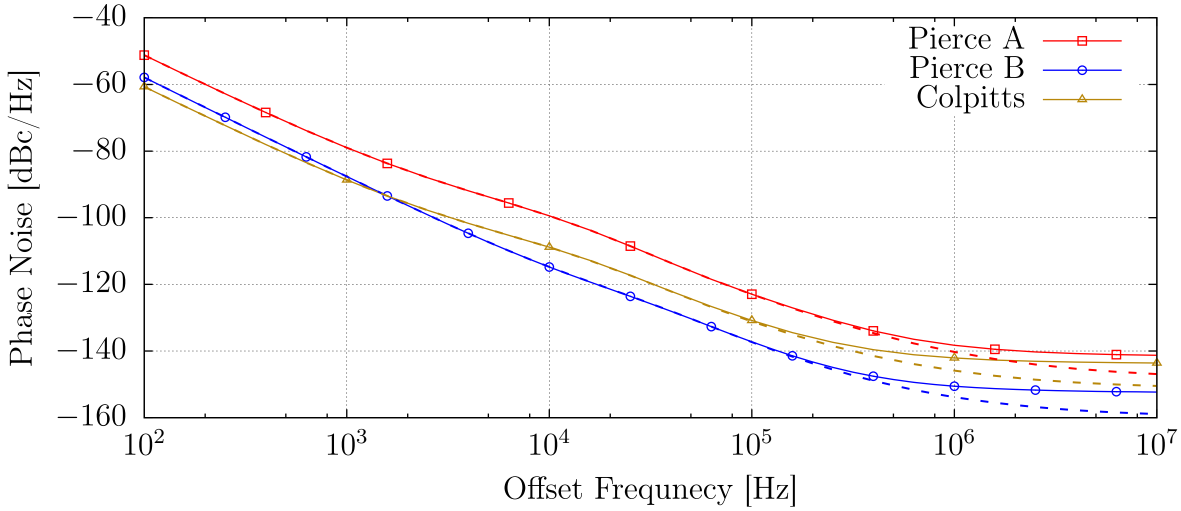


Figure 6-21: Simulated phase noise for Pierce A, Pierce B and Colpitts oscillators. Solid lines represent the noise after the output buffer while dashed lines represent the core output noise.

6.5 Oscillator Measurement Results

Both Pierce and Colpitts oscillators were fabricated in L.C. Popa's process at MTL [52]. Die micrographs for both the Pierce and Colpitts oscillators are shown in Figure 6-22 and Figure 6-23, respectively. The area of each oscillator is estimated as the smallest rectangle enclosing both the resonator and oscillator core circuit. The fabricated Pierce A, B and Collpits oscillators occupy areas of $268 \times 214\,\mu\text{m}^2$, $268 \times 205\,\mu\text{m}^2$ and $300 \times 185\,\mu\text{m}^2$, respectively.

An on-chip $50\,\Omega$ buffer with GSG output pads is included with every oscillator to drive the testing equipments without loading the oscillators. The unpassivated upper layer of large MIM capacitors was used to form DC supply pads. This is intended to provide large bypass capacitance to ground for these pads, making them reliable RF grounds and reducing the noise picked up by the DC probes.

The oscillators were tested under vacuum ($\sim 10^{-5}$ Bar) in a Cascade MicroTech PMC200 probe station [105]. Two separate Keithley 2400 SMUs were used to power the oscillator core and the $50\,\Omega$ buffer. An Agilent PXA N9030A spectrum analyzer was used to measure the oscillators' output spectrum which is shown in Figure 6-24. The output frequency of all oscillators is $\sim 1.02\,$GHz. Phase noise was also measured

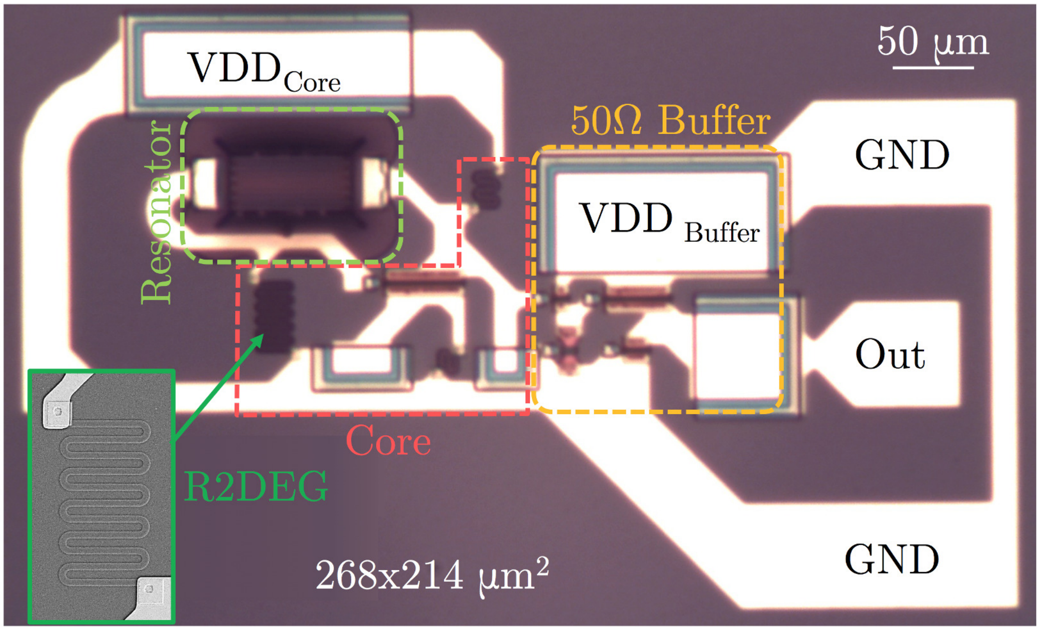


Figure 6-22: Pierce oscillator die photo showing the GaN resonator, the oscillator core including one HEMT, resistors, and capacitors along with 50 Ω buffer to drive testing equipment.

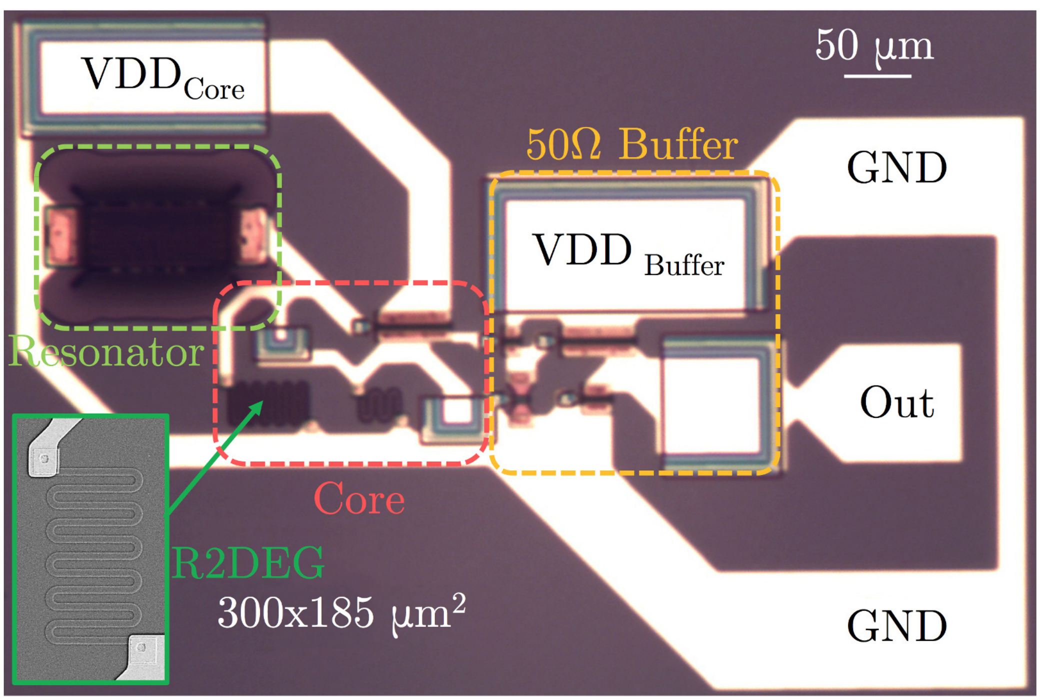


Figure 6-23: Colpitts oscillator die photo showing the GaN resonator, the oscillator core including one HEMT, resistors, and capacitors along with 50 Ω buffer to drive testing equipment.

(using the same spectrum analyzer) and is shown in Figure 6-25. At 100 kHz offset, the phase noise achieved by the oscillators is found to be −126 dBc/Hz for Pierce A, −129 dBc/Hz for Pierce B and −127 dBc/Hz for Colpitts. At 1 MHz offset, Pierce A and Colpitts achieve a phase noise of −130 dBc/Hz, whereas Pierce B achieves −141 dBc/Hz. All oscillators phase noise shows a characteristic -30 dB/dec slope corresponding to flicker ($1/f$) noise. Measured phase noise is 10 dB higher than that estimated from simulations in §6.4.3. Pierce A, B and Colpitts consume 1.06 mW, 4.7 mW and 1.47 mW, respectively, which is in exact agreement with simulations*.

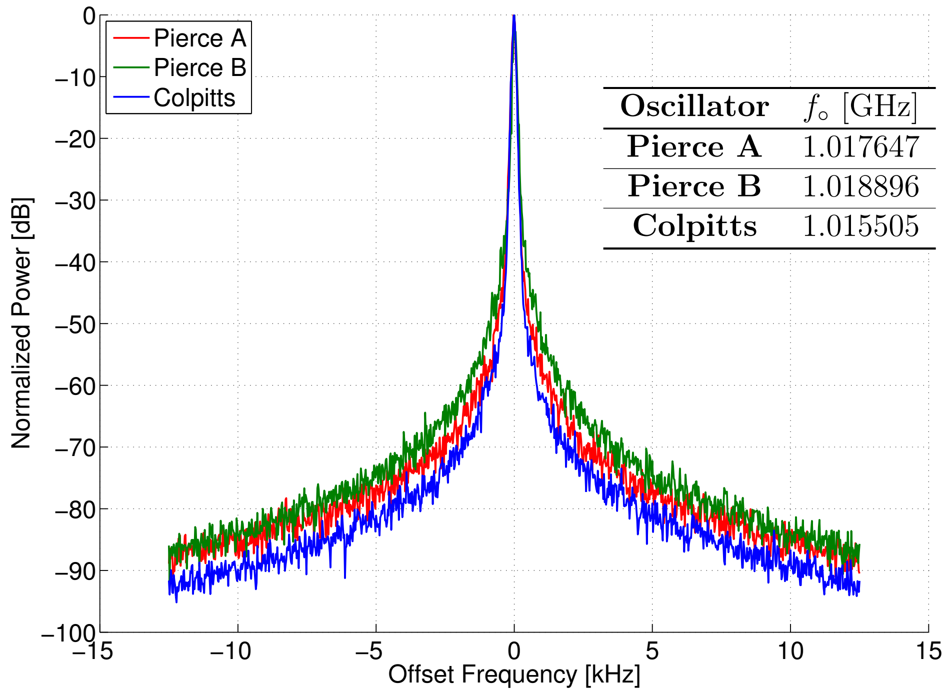
A widely accepted figure of merit (FOM) for oscillators that relates power consumption to phase noise and operating frequency is defined as

$$FOM = \left(\frac{f_\circ}{\Delta f}\right)^2 \frac{1}{\mathcal{L}(\Delta f) P_{\mathrm{DC}}}, \tag{6.28}$$

where $\mathcal{L}(\Delta f)$ and P_{DC} are the phase noise at an offset frequency Δf and the oscillator DC power consumption in milliwatts, respectively. The figure of merit (FOM) at 100 kHz is −205.9 dBc/Hz, −202.4dBc /Hz, and −205.5 dBc/Hz for Pierce A, B, and Colpitts oscillators, respectively. It is important to note that based on (6.21), operating in the Landau-Rumer regime, power consumption becomes linearly dependent on frequency, resulting in much better FOM values. In fact, in this case, a figure of merit with linear dependence on $(f_\circ/\Delta f)$ will be more appropriate.

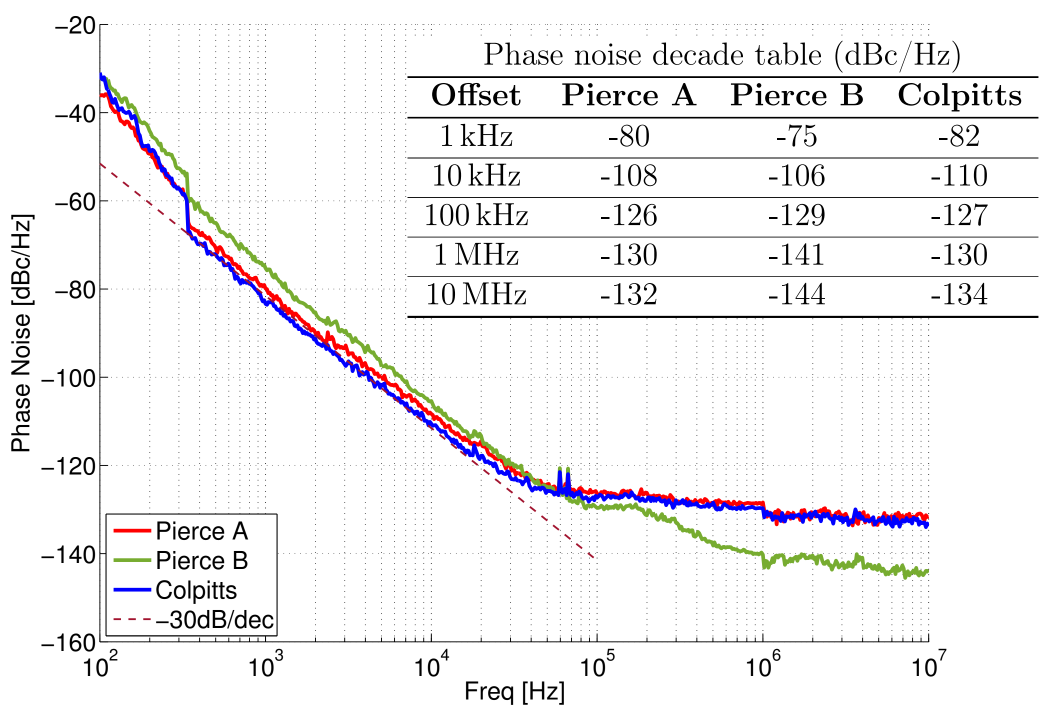
Table 6.7 on page 216 compares the oscillators implemented in this work to the state-of-the-art MEMS-based GHz-frequencies oscillators. The implemented oscillators represent the first monolithic solution with support for multiple frequency on-chip. This is attributed to the lithographically-defined frequency of the Lamb-mode resonators and the monolithic integration thereof in GaN MMICs. This provides a major advantage over thickness-mode resonators where frequency is controlled by film thicknesses. The miniature footprint together with reduced parasitics are also a direct result of such monolithic integration.

*HEMTs and passives models were fitted to test structures on the same die. DC oscillator current consumption predictions were very accurate based on these models.



Oscillator	$f_\circ$ [GHz]
Pierce A	1.017647
Pierce B	1.018896
Colpitts	1.015505

Figure 6-24: Measured output spectrum of the oscillators Pierce A, Pierce B and Colpitts.



Phase noise decade table (dBc/Hz)

Offset	Pierce A	Pierce B	Colpitts
1 kHz	-80	-75	-82
10 kHz	-108	-106	-110
100 kHz	-126	-129	-127
1 MHz	-130	-141	-130
10 MHz	-132	-144	-134

Figure 6-25: Measured phase noise of the oscillators Pierce A, Pierce B and Colpitts.

Table 6.7: Comparison between the implemented oscillators and state-of-the-art MEMS-based GHz-frequencies oscillators. This work represents the first monolithic solution with support of multi-frequency on-chip with competitive FOM and small footprint.

	This Work [52]			**CICC**	**ISSCC**	**ISSCC**	**ISSCC**
	Pierce A	**Pierce B**	**Colpitts**	**2013** [114]	**2010** [115]	**2007** [116]	**2006** [117]
Multi-$f_\circ$ on-chip	Yes			Yes	Yes	No	No
Monolithic	Yes			No	No	No	Yes
Frequency (GHz)	1.017	1.018	1.015	1.16	1.006	2.145	5.46
Core Voltage (V)	11	13	7	1	1.5	2.5	2.7
Core Current (mA)	0.096	0.362	0.21	-	-	4.8	1.7
Core Power (mW)	1.056	4.706	1.47	4.2	10.7	12	4.59
Supply Pushing (ppm/V)	2.7	1.96	1.92	-	-	65	
PN @ 100 kHz	-126	-129	-127	-144	-140*	-124	-117.7
PN@1MHz	-129	-140	-130	-173	-150	-145*	-
FOM @ 100kHz	205.9	202.4	205.5	219*	209.8*	199.8	205.8
FOM @ 1MHz	188.9	193.4	188.5	228.3	199.5*	200.8*	-
Area (mm^2)	0.057	0.055	0.056	-	0.33	0.043	0.45
Process	1 μm GaN MMIC (in-house process)			0.13 μm CMOS	0.18 μm CMOS	0.25 μm BiCOMS	0.35 μm BiCMOS
Resonator	Lamb-mode GaN			AlN CMR	Lateral Mode	BAW	FBAR

* Results are calculated or extracted from a figure, not directly provided in the corresponding reference.

6.6 Lattice and Ladder Filters

RF filters are an indispensable component of RF systems. As the integration of GaN Lamb-mode resonators in GaN MMIC technology allowed the realization of monolithic low-power oscillators; they are also suitable for implementation of both lattice and ladder filters for RF applications [150–152]. The ability of lithographically defining the resonance frequency of the Lamb-mode resonators simplifies the implementation of monolithically integrated lattice and ladder filters in this technology. This section presents the implementation and testing of lattice and ladder filters in L.C. Popa's technology.

Lattice and ladder filters are RF filters based on the electrical coupling of multiple resonators. A fully differential version of both filters is shown in Figure 6-26, where 4 resonators are required for each filter. A single-ended version of the ladder filter can also be constructed [150]. The shunt resonators in both filters have their resonance frequency $f_\circ$ smaller than the series resonators. The difference Δf between the two determines the bandwidth of the filter. The maximum attainable low-ripple fractional bandwidth is set by the resonator k^2_{eff}.

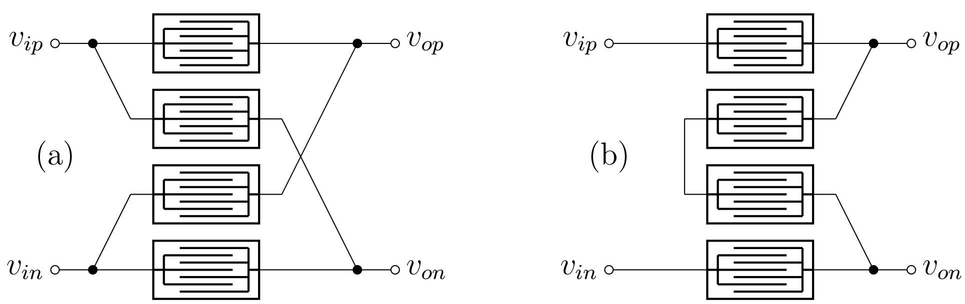


Figure 6-26: Schematic of (a) lattice filter (b) ladder filter.

Lattice filters are characterized by large out-of-band rejection, but with poor selectivity (close to carrier rejection). On the other hand, ladder filters are characterized by large selectivity and poor out-of-band rejection [150]. For this reason, both types of filters are usually combined to achieve overall good selectivity.

Both lattice and ladder filters have been implemented in L.C. Popa's process with center frequencies of 1GHz and 2GHz with different bandwidth. Figure 6-27 shows an optical micrograph of both filters, highlighting the input and output ports.

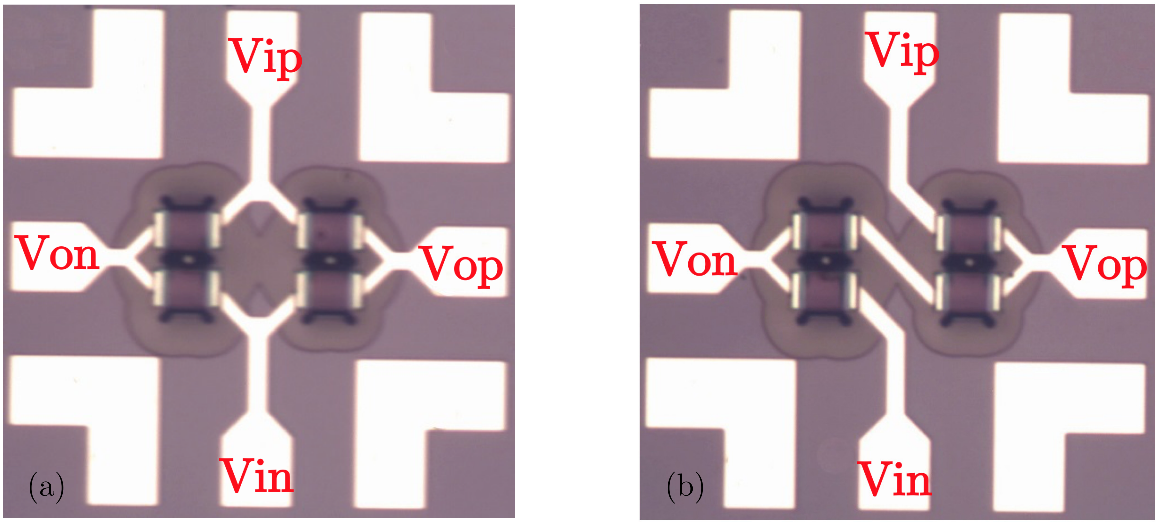


Figure 6-27: Optical micrographs of (a) lattice filters and (b) ladder filters.

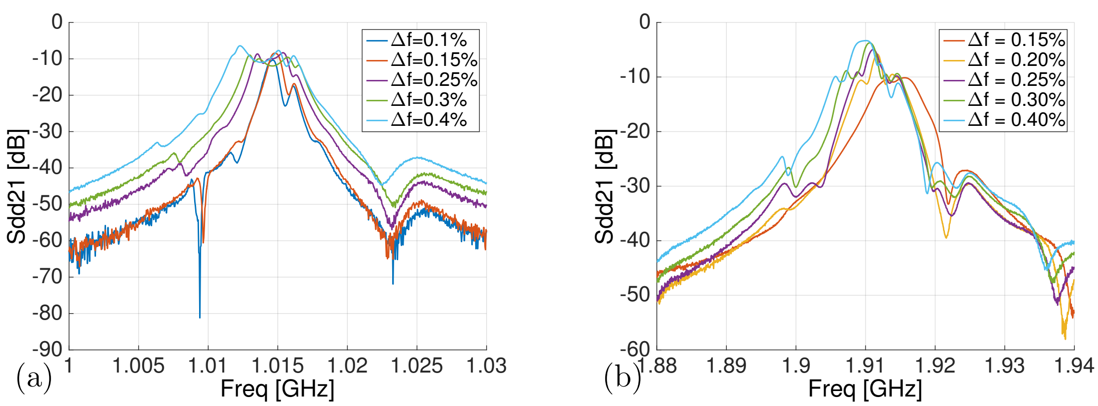


Figure 6-28: Measurement results for lattice filters designed with center frequency at (a) 1GHz and (b) 2GHz.

Full 4-port S-parameters RF measurements have been performed for all filters. Figures 6-28 and 6-29 show the insertion loss for lattice and ladder filters, respectively. The performance of the various filters demonstrated are listed in Tables* 6.8, 6.9, 6.10 and 6.11. These filters represent the first implementation of monolithically integrated MEMS-based filters in GaN MMIC technology. Together with the oscillators, they pave the road for complete monolithic RF front-ends in GaN MMICs.

*The bandwidth of some filters does not match the design Δf due to a yield issue with this particular process run.

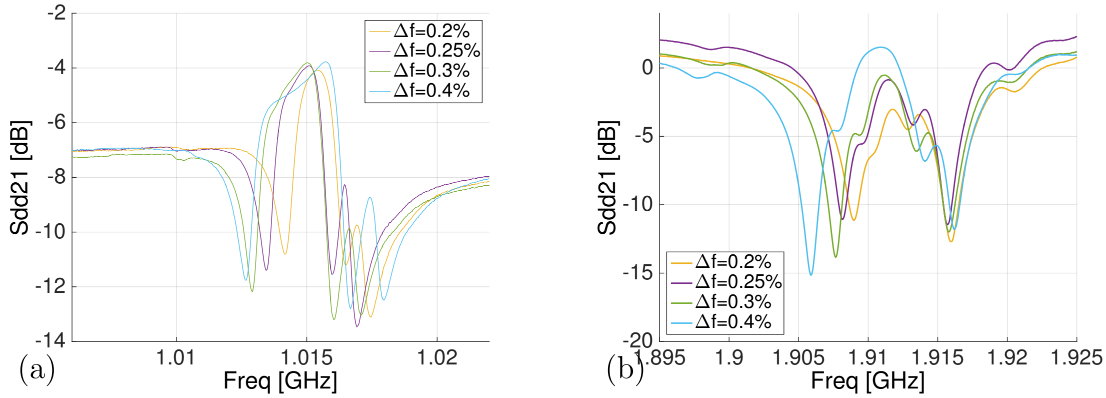


Figure 6-29: Measurement results for ladder filters designed with center frequency at (a) 1GHz and (b) 2GHz.

Table 6.8: 1 GHz lattice filters performance.

Δf [%]	$f_\circ$ [GHz]	3dB BW [MHz]	3dB BW [%]	20dB BW [MHz]	Shape Factor	IL [dB]
0.10	1.01	0.76	0.07	4.40	5.8	10.3
0.15	1.01	0.88	0.09	4.28	4.9	8.4
0.25	1.01	2.52	0.25	6.76	2.7	8.3
0.30	1.01	3.92	0.39	8.76	2.2	9.0
0.40	1.01	4.52	0.45	9.88	2.2	6.4

Table 6.9: 2 GHz lattice filters performance.

Δf [%]	$f_\circ$ [GHz]	3dB BW [MHz]	3dB BW [%]	20dB BW [MHz]	Shape Factor	IL [dB]
0.15	1.91	6.56	0.34	24.84	3.8	10.1
0.20	1.91	1.52	0.08	14.32	9.4	5.8
0.25	1.91	2.12	0.11	13.64	6.4	4.9
0.30	1.91	2.08	0.11	15.16	7.3	3.7
0.40	1.91	3.36	0.18	17.32	5.2	3.3

Table 6.10: 1 GHz ladder filters performance.

Δf [%]	$f_\circ$ [GHz]	3dB BW [MHz]	3dB BW [%]	IL [dB]
0.20	1.01	3.24	0.32	6.0
0.25	1.01	2.00	0.20	5.6
0.30	1.01	2.40	0.24	6.1
0.40	1.01	3.32	0.33	5.8

Table 6.11: 2 GHz ladder filters performance.

Δf [%]	$f_{\circ}$ [GHz]	3dB BW [MHz]	3dB BW [%]	IL [dB]
0.20	1.91	4.24	0.22	6.0
0.25	1.91	4.56	0.24	5.9
0.30	1.91	2.72	0.14	5.9
0.40	1.91	3.88	0.20	4.8

6.7 Conclusion

In conclusion, monolithic integration of GaN MEMS resonators in standard GaN MMICs has been discussed. The process developed at MIT MTL by L.C. Popa has been exploited as a potential prototyping platform, where high-Q GaN Lamb-mode MEMS resonators can be fabricated side-by-side with GaN HEMTs and other passives on the same die. The first monolithic GaN MEMS-Based oscillator has been demonstrated. Monolithic lattice and ladder filters have also been demonstrated for monolithic RF front-ends.

Extensive modeling and characterization for GaN HEMTs and other passives in this technology have been performed. The Lamb-mode resonators under consideration provided very good linearity and hence were modeled by linear equivalent circuits. HEMT models were based on the MVS model developed at MIT. The MVS model was augmented to include important physics such as g_m-dispersion and output kinks, especially useful for conservative circuit designs. A non-linear model for the meander line 2DEG resistor has also been developed based on the MVS model. Parameter extraction for all the models has been performed over multiple sets of devices.

The technological ability to integrate GaN MEMS resonators in MMIC technology, along with accurate compact models, allowed for the design, implementation, and prototyping of monolithic MEMS-based oscillators in this technology. Both Colpitts and Pierce oscillators have been designed according to classical 3-points oscillator theory, with special attention to the DC biasing requirements of the HEMTs. The designed oscillators have been fabricated at MTL by L.C. Popa, marking the first realization of monolithic MEMS-based oscillators in GaN MMIC technology.

Fabricated oscillators operate at 1 GHz with 1 mW power consumption, while occupying an area $< 0.06\,\mathrm{mm}^2$. The oscillators achieve a figure of merit around 206 dB, exceeding state of the art oscillators of comparable size and integration level. The monolithic integration of the high-Q MEMS resonator in GaN MMICs provides the unparalleled advantage of footprint and parasitics reduction, where the latter significantly contributes to the lower power consumption.

Monolithic lattice and ladder filters have also been demonstrated in the same technology. The integration of high purity oscillators and high-Q filters in GaN MMICs paves the way towards highly compact and low-power monolithic RF front-ends in GaN MMIC technologies.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 7

Conclusion and Future Directions

Towards the goal of providing ultra-compact and monolithic high-purity frequency sources and RF filters, a novel class of CMOS-MEMS resonators has been presented. These CMOS resonators are based on solid-state phononic waveguides with minimal scattering to achieve record $f_\circ \cdot Q$ products. No modifications or post-processing are required beyond the standard commercial CMOS process. Multiple RBTs were fabricated and tested in IBM 32 nm SOI technology.

Moreover, the first monolithic MEMS-based low-power high purity oscillators in standard GaN MMIC technology were presented. These oscillators are based on high performance GaN Lamb-mode resonators developed at MTL. Lattice and ladder filters have also been demonstrated in the same process.

7.1 Conclusion

Common materials in CMOS BEOL layers have been shown to present high impedance contrast and sufficient periodicity for the realization of wide bandgap PnCs in CMOS BEOL. With very simple geometry such as metal stripes, relative bandgaps as large as 85% around 4.5 GHz have been demonstrated in IBM 32 nm SOI technology. Wide bandgaps results in large reflections with a small number of PnC layers, which is common to the CMOS BEOL. The PnC geometry is similar to digital bus routing commonly found in CMOS circuits, which enhances the manufacturability of these designs. The lithographically defined dimensions of the BEOL layers were shown to be useful for PnC bandgap engineering. The effect of different CMOS process variations and mismatches on the BEOL PnCs has been studied as well. Variations only affect the edges of the bandgap, which is again mitigated by having a large bandgap.

A BEOL PnC on top of the CMOS bulk wafer has been demonstrated to form a phononic waveguide (for a particular set of wave vectors $\vec{k}$) that can trap vibrational energy in the FEOL layers. Waveguiding is achieved by virtue of the wave being unable to propagate in the PnC bandgap and by virtue of the higher sound velocity in the bulk wafer. Reflections from the bulk wafer are similar to index guiding in photonics (or incidence bigger than the critical angle in Snell's law). Although the study considered IBM 32 nm SOI technology, the SOI buried oxide (BOX) layer plays no role in elastic waveguiding, making this technique directly amenable to bulk technologies. Moreover, the dispersion characteristics of CMOS phononic waveguides can be engineered through the horizontal dimensions that can be controlled by the designer in most CMOS technologies.

The first unreleased CMOS resonant body transistor based on PnC waveguide has been demonstrated. It achieves a quality factor of 252 at 2.81 GHz, marking an 8$\times$ improvement in Q over previous RBTs. Single-ended drive and sense were used for this first PnC-Based RBT.

Next, coupled mode theory for both non-grated and grated waveguides have been explored, in explicit analogy to photonics waveguide and quantum mechanics. Fully differential driving and slowly varying transitions are demonstrated to be two techniques to reduce scattering in unreleased MEMS resonators. The former guarantees the operation as far as possible from the sound cone, whereas the latter follow from the adiabatic theorem. The advantages of reducing scattering in the resonator are an increased quality factor as well as the reduction of spurious modes.

The aforementioned phononic waveguides have then been demonstrated to implement high-Q unreleased CMOS RBTs. Fully differential driving and sensing are used to guarantee waveguiding, resulting in a balanced-balanced device. Electrical isolation between drive and sense is also necessary, yet careful consideration is required for the structure to match the different sections of the waveguide. Methods for horizontal energy confinement in the phononic waveguide RBTs are also elaborated. Using different waveguide sections with mismatched dispersion characteristics proves to be a good choice in terms of reducing scattering. Adiabatic transition to termination can also help with scattering reduction and provides a better stress distribution for the main cavity waveguide.

Numerical optimization for CMOS RBTs is necessary to optimize the quality factor as well as the output signal amplitude. Matching the different sections of the cavity is an implicit requirement in such optimization problems. A numerical framework based on model order reduction and memoization is proposed for up to 40× speed-up in full structural optimization of the RBT. The optimization flow allowed for optimizing the RBTs under consideration and can be directly applied to any MEMS device that relies on frequency domain FEM simulation for determining its performance.

Optimized RBTs have been fabricated in IBM 32 nm SOI technology without any process modifications or post-processing. RBTs based on phononic waveguides show superior energy confinement. This results in fabricated devices with 58× improvement in Q and 68× in $f_\circ \cdot Q$ over previous generation RBTs. The presented CMOS RBTs have the highest Q and $f_\circ \cdot Q$ to date. It is also a solid indication that the previous generations were not limited by intrinsic material losses.

Moreover, RBTs in FinFET CMOS technologies have been explored. The fins' small feature size allows for frequency scaling up to tens of gigahertz; while their structure provide superior acoustic confinement and strong coupling to the driving stress. High Q and larger output signals are the immediate benefits. FinFET RBTs, with resonance frequencies around 33 GHz and 10 GHz, have been implemented and tapedout in Globalfoundries 14 nm technology. Such RBTs can enable unprecedented monolithic low-phase noise, low-power oscillators at such high frequencies. Those oscillators can be used for RF applications, or to build large coupled clusters for exceedingly fast unconventional signals processing.

Also, a practical approach towards RBT compact modeling has been presented, which allows for efficient simulation of the RBT tightly coupled physics, without the need for computationally intensive FEM field solutions. The presented models are implemented in Verilog-A and are fully compatible with commercial EDA software. A modular approach has been adopted, where individual physics are modeled by separate modules interacting through a set of nodes. This allows for seamless model expansion and development.

Equivalent circuit and transmission lines were used to model mechanical resonance, with the latter providing more accurate parameterization. Parallel-plate capacitor and MOSCAP transducers have also been modeled, with special attention to numerical hazards. FET sensing based on the full BSIM and MVS transistor model has been implemented. An equivalent circuit model has been adopted for self-heating, whereas a full parasitics network is also included. These compact models allow circuits and systems designers to seamlessly integrate them into their workflow.

Monolithic integration of GaN MEMS resonators in standard GaN MMICs has also been explored within the prototyping process developed at MIT MTL by L.C. Popa. In this technology, High-Q GaN Lamb-mode MEMS resonators can be fabricated side-by-side with GaN HEMTs and other passives on the same die. The first monolithic GaN MEMS-Based oscillator has been demonstrated together with monolithic lattice and ladder filters for complete monolithic RF front-ends in GaN MMICs technology.

Extensive modeling, characterization and parameters extraction for GaN HEMTs and other passives in this technology have been performed. The Lamb-mode resonators under consideration demonstrated very good linearity and hence were modeled by linear equivalent circuits. HEMT models were based on the MVS model developed at MIT, augmented to include g_m dispersion and output kinks. A non-linear model for meander line 2DEG resistors has also been developed based on the MVS model.

The first realization of monolithic MEMS-based oscillators in GaN MMIC technology is demonstrated thanks to the technological ability to integrate GaN MEMS resonators in MMIC technology along with accurate compact models. Both Colpitts and Pierce oscillators have been designed according to classical 3-points oscillator theory, with special attention to the HEMTs DC biasing requirements.

The designed and laid-out oscillators have been fabricated at MTL by L.C. Popa. They operate at 1 GHz with 1 mW power consumption, while occupying an area $< 0.06\,\mathrm{mm}^2$. The oscillators achieve a figure of merit around 206 dB, exceeding state of the art oscillators of comparable size and integration level. The monolithic integration of the high-Q MEMS resonator in GaN MMICs provides the unparalleled advantage of footprint and parasitics reduction, where the latter significantly contributes to the lower power consumption. Monolithic lattice and ladder filters have also been demonstrated in the same technology, allowing for complete low-power monolithic RF front-ends in GaN MMIC technology.

The monolithic integration of MEMS resonators, high purity oscillators and high-Q filters in standard IC technology is presented as a potential solution for the ever expanding demands of compact, low-power, and high-frequency RF communication systems.

7.2 Future Directions

7.2.1 Thermal Stability and Packaging

The presented CMOS RBTs are expected to show a small temperature coefficient of frequency (TCF). This is due to the fact that the materials forming the cavity have Young's modulus temperature coefficients (TCE) with opposite signs [36]. Such materials combination leads to passive temperature compensation as discussed in [70]. The direct result is a better TCF as compared to single material resonators. A natural extension of the work presented in this thesis is to study the temperature stability of the CMOS RBTs based on phononic waveguides.

It is also important to study the effects of packaging on the resonator performance. Packaging-induced stress results due to the difference in temperature expansion coefficients between the CMOS die and package material. This creates strain in the FEOL layers of the CMOS die, effectively changing the resonator dimensions. The work presented in this thesis can be further augmented to study the effects of package-induced stress on the presented CMOS RBTs.

7.2.2 Advanced Phononic Circuits and Signal Processing

In chapter 3, an operator-theoretic analytical framework has been developed for elastic wave propagation and phononic waveguides. In this development, explicit analogy to quantum mechanics and photonic systems has been presented. This establishes a unifying theoretical framework for the treatment of systems admitting generalized Hermitian formulations. A major future direction would be to investigate importing further theoretical results from the quantum mechanics and photonics community to phononic structures.

In the light of these theoretical developments, many well-established concepts and structures, particularly in the photonics community, can be readily applied to CMOS-integrated phononic waveguides. On-chip delay lines, couplers, matching networks, and antennas are all good examples for potential on-chip phononic components. This allows for advanced on-chip electromechanical phononic circuits useful for a multitude of applications from advanced signal processing to studying macroscopic quantum mechanics [153]. Studying and prototyping these potential components and systems is a natural extension for the waveguide study presented in this work.

7.2.3 Ferroelectric Transduction

As observed in §4.5, the electromechanical transconductance of CMOS RBTs is dropping below 100 nS. With such low signal levels, the implementation of filters and oscillators is challenging. It is thus instructive to investigate possible solutions to enhance the RBT signal level.

Optimization of the mechanical structure and quality factor may help improve the output signal. However, the demonstrated resonators achieve very high quality factors, and further structural optimization is not expected to yield the required improvements. Different transduction physics are required to boost the resonator signal.

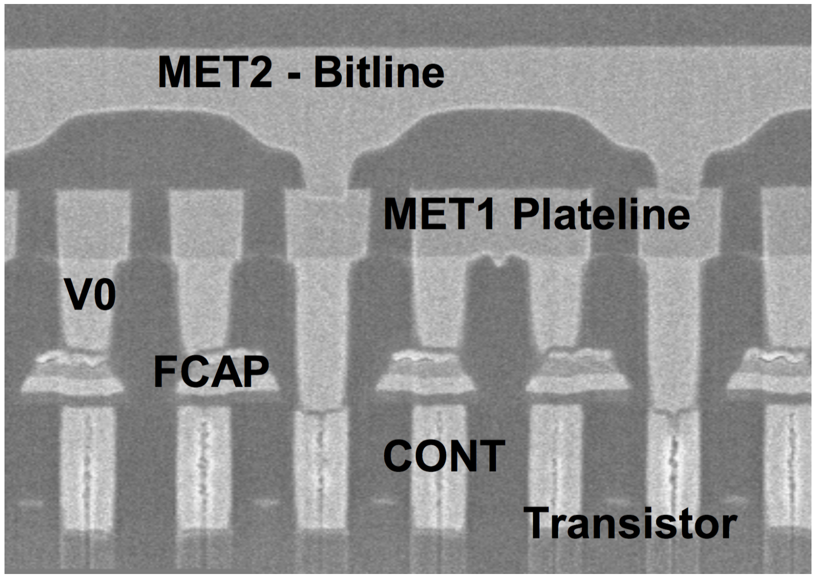


Figure 7-1: FeCAPs in Texas Instruments CMOS process after [154].

Piezoelectric transduction is a potential solution for improving the RBT signal amplitude. When compared to electrostatic transduction, piezoelectric transduction is usually characterized by orders of magnitude better electromechanical transduction efficiency. Piezoelectric materials have been recently introduced into CMOS technologies in the form of ferroelectric materials [154–156]. These are generally used for the ferroelectric RAM (FeRAM) memories. Figure 7-1 shows an early implementation of standalone ferroelectric capacitors (FeCAPs) in Texas Instruments' CMOS process after [154]. Another implementation uses ferroelectric HfO_2 as the gate dielectric in a high-κ metal gate (HKMG) transistor, as in IPMS-CNT and Globalfoundries process [155, 156]. This gate stack is shown in Figure 7-2.

Both implementations can be used for driving and sensing unreleased CMOS resonators. The ferroelectric HKMG has an advantage when used as sensing transistor in CMOS-RBT. Basically, the strain in the transistor causes modulation of the channel charge as opposed to the channel mobility. This provides a much larger signal as it relies on the transistor electrical g_m. It is also suitable for aggressively scaled ballistic FETs, where the current is insensitive to the actual channel mobility.

An important extension of the work presented in this thesis would be a comprehensive study of the use of CMOS FeRAM devices to boost the RBT signal amplitude.

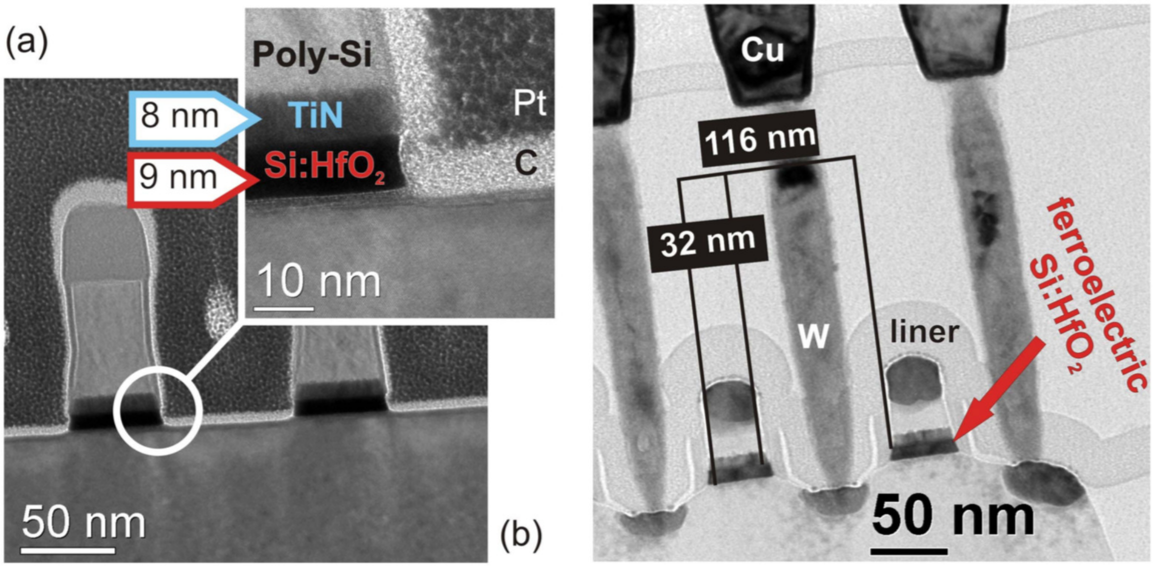


Figure 7-2: Ferroelectric HfO_2 gate oxide for HKMG FeRAM transistor after [155].

This improvement is required to lower the insertion loss of the BRT, significantly reducing the required amplifier gain to close the oscillator loop and enabling on-chip low-power, low-phase noise miniaturized oscillators. These RBT-based oscillators can be colloquially referred to as resonant body oscillators or RBOs.

7.2.4 Unconventional Signal Processing and Cognitive Memory

The monolithic integration and miniaturization of low-power, low-phase-noise oscillators enable the realization of large on-chip arrays of coupled oscillators. Synchronization dynamics in these oscillator arrays are strong function of coupling configurations. Many applications are based on such dynamics and the control thereof. Phase noise reduction, cognitive memories and unconventional processing of signals and data exploitation (UPSDIE) are few examples.

The author has previously participated in multiple theoretical studies of these systems [157–160]. Further development and prototyping of such systems based on RBOs is a potential extension for the work in this thesis.

7.2.5 CMOS Opto-mechanics

Silicon photonics have seen a major development over the past decade. Photonic waveguides, modulators and photodetectors have all been demonstrated in commercial CMOS SOI technology by Prof. R. Ram's Group at MIT [161]. In particular, photonic waveguides are implemented from the gate/active area stack of the CMOS SOI technology. For the unreleased CMOS phononic waveguides and MEMS resonators presented in this work, the strain-energy and peak stress are also confined around the same gate stack.

This has the potential of creating photonic-phononic waveguides with photon-phonon coupling. Such waveguides could enable a multitude of applications such as the realization of RF-photonic signal processing, narrow-line-width laser sources, and RF-waveform synthesis [162]. This is also an inviting opportunity to investigate cavity opto-mechanics [163] with its countless applications and the resulting benefits from monolithic integration in CMOS technologies.

7.2.6 GaN MMICs: RF Front-Ends and Beam Forming

In this thesis, the first low-power, low-phase-noise, monolithic MEMS-based oscillator in GaN has been demonstrated, along with lattice and ladder filters. The demonstration of these components prove that the realization of complete low-power RF front-ends in GaN MMICs is quite feasible. Prototyping such front-ends is a natural extension to the work presented in this thesis.

Moreover, arrays of coupled oscillators can also be integrated in GaN MMICs. Such arrays can be useful for phase noise reduction, UPSIDE or synthesizing multi-phase signals. Such possibilities combined with the high-power capabilities of GaN MMIC technology could be revolutionary for beam forming and radar applications. Further investigation of these topics is proposed.

THIS PAGE INTENTIONALLY LEFT BLANK

Appendix A

Hermiticity Proof of $\hat{A}$

In this appendix we prove the Hermiticity of the operator $\hat{A}$ on the left-hand-side of (3.11a). First, following the proof in [70], consider

$$\left\langle \vec{v}' \middle| \nabla_{tr} \cdot \vec{T} \right\rangle_{\partial\Omega} = \int_{\partial\Omega} {v'_x}^* \left(\partial_x T_{xx} + \partial_y T_{xy}\right) + \int_{\partial\Omega} {v'_y}^* \left(\partial_x T_{xy} + \partial_y T_{yy}\right) + \int_{\partial\Omega} {v'_z}^* \left(\partial_x T_{xz} + \partial_y T_{yz}\right). \tag{A.1}$$

Using the vector identity

$$\boldsymbol{\nabla} \cdot \left(\psi \vec{A}\right) = \psi \boldsymbol{\nabla} \cdot \vec{A} + \boldsymbol{\nabla} \psi \cdot \vec{A}, \tag{A.2}$$

we can re-write (A.1) as

$$\begin{aligned} \left\langle \vec{v}' \middle| \nabla_{tr} \cdot \vec{T} \right\rangle_{\partial\Omega} = & \int_{\partial\Omega} \boldsymbol{\nabla} \cdot \left({v'_x}^* \begin{bmatrix} T_{xx} \\ T_{xy} \end{bmatrix} \right) - \int_{\partial\Omega} \boldsymbol{\nabla} \left({v'_x}^*\right) \cdot \begin{bmatrix} T_{xx} \\ T_{xy} \end{bmatrix} \\ & + \int_{\partial\Omega} \boldsymbol{\nabla} \cdot \left({v'_y}^* \begin{bmatrix} T_{xy} \\ T_{yy} \end{bmatrix} \right) - \int_{\partial\Omega} \boldsymbol{\nabla} \left({v'_y}^*\right) \cdot \begin{bmatrix} T_{xy} \\ T_{yy} \end{bmatrix} \\ & + \int_{\partial\Omega} \boldsymbol{\nabla} \cdot \left({v'_z}^* \begin{bmatrix} T_{xz} \\ T_{yz} \end{bmatrix} \right) - \int_{\partial\Omega} \boldsymbol{\nabla} \left({v'_z}^*\right) \cdot \begin{bmatrix} T_{xz} \\ T_{yz} \end{bmatrix}. \end{aligned} \tag{A.3}$$

By using a 2D version of the divergence theorem, all the divergence terms become an integration over a path $\partial^2\Omega$, surrounding the cross-sectional area $\partial\Omega$ of the waveguide. For guided (confined) waveguide modes, these path integrals vanish. This can be justified by considering the integration domain $\partial\Omega$ to extend indefinitely, where the stress and velocity fields vanish along the surrounding path $\partial^2\Omega$. Thus, (A.3) simplifies to

$$\begin{aligned}\left\langle \vec{v}\,' \middle| \nabla_{tr} \cdot \vec{T} \right\rangle_{\partial\Omega} = -\int_{\partial\Omega} &(\partial_x {v'_x}^*) T_{xx} + (\partial_y {v'_x}^*) T_{xy} + (\partial_x {v'_y}^*) T_{xy} \\ &+ (\partial_y {v'_y}^*) T_{yy} + (\partial_x {v'_z}^*) T_{xz} + (\partial_y {v'_z}^*) T_{yz}, \end{aligned} \tag{A.4}$$

from which we conclude that

$$\left\langle \vec{v}\,' \middle| \nabla_{tr} \cdot \vec{T} \right\rangle_{\partial\Omega} = -\left\langle \nabla_{tr-s} \vec{v}\,' \middle| \vec{T} \right\rangle_{\partial\Omega} \tag{A.5}$$

and hence $(i\nabla_{tr-s})$ is the adjoint of $(i\nabla_{tr}\cdot)$. Moreover, for lossless materials, the compliance matrix s is a real and symmetric matrix (small losses can be added by perturbation theory). This makes the operator $\hat{A}$ in (3.11a) Hermitian.

Appendix B

Piezoresistivity Coefficients

B.1 Piezoresistivity Tensors and Coordinate Transformation

The material resistivity tensor ρ_{ij} changes linearly with the stress components T_{ij} as [164]

$$\Delta\rho_{ij} = \rho_{ij} - \rho_{ij}^{0} = \bar{\rho}\pi_{ijkl}T_{kl}, \tag{B.1}$$

where π_{ijkl} are the components of the fourth-rank piezoresistivity tensor, with the mean unstressed resistivity given by

$$\bar{\rho} = \frac{1}{3}\left(\rho_{xx}^{0} + \rho_{yy}^{0} + \rho_{zz}^{0}\right). \tag{B.2}$$

The piezoresistivity tensor is symmetric such that

$$\pi_{ijkl} = \pi_{jikl} = \pi_{ijlk}. \tag{B.3}$$

This allows the use of the abbreviated notation

$$\Delta\rho_{\alpha} = \bar{\rho}\pi_{\alpha\beta}T_{\beta}, \tag{B.4}$$

	$\rho(\Omega cm)$	π_{11}	π_{12}	π_{44}
n-type $\langle 100 \rangle$ Si	11.7	- 102.2	53.4	-13.6
p-type $\langle 100 \rangle$ Si	7.8	6.6	-1.1	138.1

Table B.1: Piezoresistive coefficients for $\langle 100 \rangle$ silicon in units of $10^{-11} Pa^{-1}$.

with the indices $\alpha,\beta = 1,2,3,4,5,6$. The stresses in abbreviated notation correspond to

$$\begin{aligned} T_1 &= T_{xx} \qquad T_2 = T_{yy} \qquad T_3 = T_{zz} \\ T_4 &= T_{yz} \qquad T_5 = T_{zx} \qquad T_6 = T_{xy}. \end{aligned} \tag{B.5}$$

The resistivity change $\Delta\rho_\alpha$ follows the same index convention. The factors of 2 resulting from the use of the abbreviated subscripts are considered into the coefficients $\pi_{\alpha\beta}$. For cubic crystal, the change in resistivity is thus formulated as

$$\boldsymbol{\Delta\rho} = \bar{\rho}\boldsymbol{\pi}\boldsymbol{T} \tag{B.6}$$

with

$$\boldsymbol{\Delta\rho} = \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \\ \rho_5 \\ \rho_6 \end{pmatrix}, \ \boldsymbol{T} = \begin{pmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \\ T_5 \\ T_6 \end{pmatrix} \text{ and } \boldsymbol{\pi} = \begin{pmatrix} \pi_{11} & \pi_{12} & \pi_{12} & 0 & 0 & 0 \\ \pi_{12} & \pi_{11} & \pi_{12} & 0 & 0 & 0 \\ \pi_{12} & \pi_{12} & \pi_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & \pi_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \pi_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \pi_{44} \end{pmatrix}. \tag{B.7}$$

The piezoresistive coefficients for $\langle 100 \rangle$ silicon are listed in table B.1.

For RBTs fabricated in CMOS technology, the sensing FET channel is usually aligned with $\langle 110 \rangle$ direction. It becomes more convenient to adopt a rotated coordinate system along crystallographic directions $\langle 110 \rangle$, $\langle \bar{1}10 \rangle$ and $\langle 100 \rangle$. The second-rank tensors transformation matrix becomes in this case

$$\boldsymbol{a} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{B.8}$$

By using this second-rank tensor transformation, the fourth-rank transformation matrix with abbreviated subscripts can be easily obtained as described in [68] and [164]. The transformed piezoresistivity tensor is given by

$$\boldsymbol{\pi'} = \begin{pmatrix} \boldsymbol{\pi'}_I & 0 \\ 0 & \boldsymbol{\pi'}_{II} \end{pmatrix}, \tag{B.9}$$

with

$$\boldsymbol{\pi'}_I = \begin{pmatrix} \frac{\pi_{11}+\pi_{12}+\pi_{44}}{2} & \frac{\pi_{11}+\pi_{12}-\pi_{44}}{2} & \pi_{12} \\ \frac{\pi_{11}+\pi_{12}-\pi_{44}}{2} & \frac{\pi_{11}+\pi_{12}+\pi_{44}}{2} & \pi_{12} \\ \pi_{12} & \pi_{12} & \pi_{11} \end{pmatrix}, \; \boldsymbol{\pi'}_{II} = \begin{pmatrix} \pi_{44} & 0 & 0 \\ 0 & \pi_{44} & 0 \\ 0 & 0 & \pi_{11}-\pi_{12} \end{pmatrix}.$$

B.2 FET Channel as Two-Terminals Piezoresistor

Estimating the change in resistance for a two-terminal piezoresistor is essential in determining the change in the mobility of the sensing FET channel. For a two-terminal piezoresistor of length l along the direction $\vec{n}$, the voltage drop V is given by

$$V = \vec{E}.\vec{n}l, \tag{B.10}$$

where $\vec{E}$ is the electric field. The current density is related to the electric field through the relation

$$E_i = \rho_{ij}J_j = \bar{\rho}(\delta_{ij} + \pi'_{ijkl}T'_{kl})J_j. \tag{B.11}$$

The resistance of the element is then defined as

$$R = \frac{V}{I} = \frac{l\vec{E}.\vec{n}}{A\vec{J}} = \frac{l\vec{E}.\vec{n} \otimes \vec{J}}{A\vec{J} \otimes \vec{J}}, \tag{B.12}$$

where $\otimes$ represents the dyadic tensor product. Substituting $\vec{E} = E_i n_i$ and $\vec{J} = J n_j$, with $\sqrt{n_i n_i} = 1$ into (B.12), one gets

$$R = \frac{\bar{\rho}l}{A}(\delta_{ij} + \pi'_{ijkl}T'_{kl})n_i n_j. \tag{B.13}$$

The unstressed resistance is given by

$$R^0 = \frac{\bar{\rho} l}{A}(\delta_{ij}) n_i n_j = \frac{\bar{\rho} l}{A}. \tag{B.14}$$

The relative change in the resistance of the element is then given by

$$\frac{\Delta R}{R} = \frac{R - R^0}{R^0} = n_i n_j \pi'_{ijkl} T'_{kl}. \tag{B.15}$$

Considering the FET channel to be along crystallographic direction $\langle 110 \rangle$, the x-axis of the rotated coordinate system, we have $n_x = 1$ while $n_y = n_z = 0$. Equation (B.15) reduces to

$$\frac{\Delta R}{R} = \pi'_{1\beta} T'_{\beta}. \tag{B.16}$$

With the coefficients given in (B.9), the change in resistance is given by

$$\frac{\Delta R}{R} = \frac{1}{2}(\pi_{11} + \pi_{12} + \pi_{44}) T_1 + \frac{1}{2}(\pi_{11} + \pi_{12} - \pi_{44}) T_2 + \pi_{12} T_3. \tag{B.17}$$

For the 2D simulation approximation used in COMSOL, T_1, T_2 and T_3 correspond to T_{xx}, T_{zz} and T_{yy}, respectively. The change in the FET channel mobility can be directly estimated from COMSOL simulation results as:

$$\frac{\Delta \mu}{\mu} = -\frac{\Delta R}{R} = -\frac{1}{2}(\pi_{11} + \pi_{12} + \pi_{44}) T_{xx} - \frac{1}{2}(\pi_{11} + \pi_{12} - \pi_{44}) T_{zz} - \pi_{12} T_{yy}. \tag{B.18}$$

It is important to note that with plane strain approximation, the out-of-plane stress component T_{zz} is smaller than T_{xx} and T_{yy}, yet it is not trivial.

Bibliography

[1] Joydeep Basu and Tarun Kanti Bhattacharyya. Microelectromechanical resonators for radio frequency communication applications. *Microsystem Technologies*, 17(10):1557–1580, 2011. URL: `http://dx.doi.org/10.1007/s00542-011-1332-9`, `doi:10.1007/s00542-011-1332-9`.

[2] D.T. Chang and R.L. Kubena. Large area integration of quartz resonators with electronics, Feb 2008. US Patent App. 11/502,336. URL: `http://www.google.ch/patents/US20080034575`.

[3] C. T. c. Nguyen. MEMS technology for timing and frequency control. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 54(2):251–270, February 2007. `doi:10.1109/TUFFC.2007.240`.

[4] C. S. Lam. A review of the recent development of MEMS and crystal oscillators and their impacts on the frequency control products industry. In *Ultrasonics Symposium, 2008. IUS 2008. IEEE*, pages 694–704, Nov 2008. `doi:10.1109/ULTSYM.2008.0167`.

[5] A. Partridge, Hae-Chang Lee, P. Hagelin, and V. Menon. We know that MEMS is replacing quartz. But why? And why now? In *European Frequency and Time Forum International Frequency Control Symposium (EFTF/IFC), 2013 Joint*, pages 411–416, July 2013. `doi:10.1109/EFTF-IFC.2013.6702311`.

[6] G.K. Fedder, R.T. Howe, Tsu-Jae King Liu, and E.P. Quevy. Technologies for Cofabricating MEMS and Electronics. *Proceedings of the IEEE*, 96(2):306–322, Feb 2008. `doi:10.1109/JPROC.2007.911064`.

[7] H.C. Nathanson, W.E. Newell, R.A. Wickstrom, and Jr. Davis, J.R. The resonant gate transistor. *Electron Devices, IEEE Transactions on*, 14(3):117–133, Mar 1967. `doi:10.1109/T-ED.1967.15912`.

[8] W. C. Tang, T. C. H. Nguyen, and R. T. Howe. Laterally driven polysilicon resonant microstructures. In *Micro Electro Mechanical Systems, 1989, Proceedings, An Investigation of Micro Structures, Sensors, Actuators, Machines and Robots. IEEE*, pages 53–59, February 1989. `doi:10.1109/MEMSYS.1989.77961`.

[9] Kun Wang, Ark-Chew Wong, and C. T. C. Nguyen. VHF free-free beam high-Q micromechanical resonators. *Journal of Microelectromechanical Systems*, 9(3):347–360, Sept 2000. `doi:10.1109/84.870061`.

[10] Yu-Wei Lin, Seungbae Lee, Sheng-Shian Li, Yuan Xie, Zeying Ren, and C. T. Nguyen. Series-resonant VHF micromechanical resonator reference oscillators. *IEEE Journal of Solid-State Circuits*, 39(12):2477–2491, Dec 2004. doi:10.1109/JSSC.2004.837086.

[11] S. Pourkamali, A. Hashimura, R. Abdolvand, G. K. Ho, A. Erbil, and F. Ayazi. High-Q single crystal silicon HARPSS capacitive beam resonators with self-aligned sub-100-nm transduction gaps. *Journal of Microelectromechanical Systems*, 12(4):487–496, Aug 2003. doi:10.1109/JMEMS.2003.811726.

[12] S. A. Bhave, Di Gao, R. Maboudian, and R. T. Howe. Fully-differential poly-SiC Lame mode resonator and checkerboard filter. In *Micro Electro Mechanical Systems, 2005. MEMS 2005. 18th IEEE International Conference on*, pages 223–226, Jan 2005. doi:10.1109/MEMSYS.2005.1453907.

[13] L.C. Popa and D. Weinstein. L-band Lamb mode resonators in Gallium Nitride MMIC technology. In *Frequency Control Symposium (FCS), 2014 IEEE International*, pages 1–4, May 2014. doi:10.1109/FCS.2014.6859997.

[14] R. Ruby, P. Bradley, III Larson, J., Y. Oshmyansky, and D. Figueredo. Ultra-miniature high-Q filters and duplexers using FBAR technology. In *Solid-State Circuits Conference, 2001. Digest of Technical Papers. ISSCC. 2001 IEEE International*, pages 120–121, Feb 2001. doi:10.1109/ISSCC.2001.912569.

[15] R. Aigner, J. Ella, H.-J. Timme, L. Elbrecht, W. Nessler, and S. Marksteiner. Advancement of MEMS into RF-filter applications. In *Electron Devices Meeting, 2002. IEDM '02. International*, pages 897–900, Dec 2002. doi:10.1109/IEDM.2002.1175981.

[16] R. Aigner. High performance RF-filters suitable for above IC integration: film bulk-acoustic- resonators (FBAR) on silicon. In *Custom Integrated Circuits Conference, 2003. Proceedings of the IEEE 2003*, pages 141–146, Sept 2003. doi:10.1109/CICC.2003.1249378.

[17] S.D. Senturia. *Microsystem Design*. Springer US, 2005.

[18] C.B. O'Neal, AP. Malshe, S.B. Singh, W.D. Brown, and W.P. Eaton. Challenges in the packaging of MEMS. In *Advanced Packaging Materials: Processes, Properties and Interfaces, 1999. Proceedings. International Symposium on*, pages 41–47, Mar 1999. doi:10.1109/ISAPM.1999.757284.

[19] K. Persson and K. Boustedt. Fundamental requirements on MEMS packaging and reliability. In *Advanced Packaging Materials, 2002. Proceedings. 2002 8th International Symposium on*, pages 1–7, 2002. doi:10.1109/ISAPM.2002.990357.

[20] A Morris and S. Cunningham. Challenges and Solutions for Cost-effective RF-MEMS Packaging. In *Electronic Manufacturing Technology Symposium, 2007.*

IEMT '07. 32nd IEEE/CPMT International, pages 278–285, Oct 2007. `doi:10.1109/IEMT.2007.4417077`.

[21] Huikai Xie, G.K. Fedder, et al. Post-CMOS processing for high-aspect-ratio integrated silicon microstructures. *Microelectromechanical Systems, Journal of*, 11(2):93–101, Apr 2002. `doi:10.1109/84.993443`.

[22] Cheng-Syun Li, Li-Jen Hou, and Sheng-Shian Li. Advanced CMOS-MEMS Resonator Platform. *Electron Device Letters, IEEE*, 33(2):272–274, Feb 2012. `doi:10.1109/LED.2011.2175695`.

[23] C. S. Sander, J. W. Knutti, and J. D. Meindl. A monolithic capacitive pressure sensor with pulse-period output. *IEEE Transactions on Electron Devices*, 27(5):927–930, May 1980. `doi:10.1109/T-ED.1980.19958`.

[24] J. H. Smith, S. Montague, J. J. Sniegowski, J. R. Murray, and P. J. McWhorter. Embedded micromechanical devices for the monolithic integration of MEMS with CMOS. In *Electron Devices Meeting, 1995. IEDM '95., International*, pages 609–612, Dec 1995. `doi:10.1109/IEDM.1995.499295`.

[25] R. N. Candler, Woo-Tae Park, Huimou Li, G. Yama, A. Partridge, M. Lutz, and T. W. Kenny. Single wafer encapsulation of MEMS devices. *IEEE Transactions on Advanced Packaging*, 26(3):227–232, Aug 2003. `doi:10.1109/TADVP.2003.818062`.

[26] L. J. Hornbeck. Digital Light Processing and MEMS: an overview. In *Advanced Applications of Lasers in Materials Processing/Broadband Optical Networks/Smart Pixels/Optical MEMs and Their Applications. IEEE/LEOS 1996 Summer Topical Meetings:*, pages 7–8, Aug 1996. `doi:10.1109/LEOSST.1996.540770`.

[27] Qing Zhang, P. V. Cicek, K. Allidina, F. Nabki, and M. N. El-Gamal. Surface-Micromachined CMUT Using Low-Temperature Deposited Silicon Carbide Membranes for Above-IC Integration. *Journal of Microelectromechanical Systems*, 23(2):482–493, April 2014. `doi:10.1109/JMEMS.2013.2281304`.

[28] F. Nabki, T. A. Dusatko, S. Vengallatore, and M. N. El-Gamal. Low-Stress CMOS-Compatible Silicon Carbide Surface-Micromachining Technology;Part I: Process Development and Characterization. *Journal of Microelectromechanical Systems*, 20(3):720–729, June 2011. `doi:10.1109/JMEMS.2011.2111355`.

[29] F. Nabki, K. Allidina, F. Ahmad, P. V. Cicek, and M. N. El-Gamal. A Highly Integrated 1.8 GHz Frequency Synthesizer Based on a MEMS Resonator. *IEEE Journal of Solid-State Circuits*, 44(8):2154–2168, Aug 2009. `doi:10.1109/JSSC.2009.2022914`.

[30] Wen-Chien Chen, Weileun Fang, and Sheng-Shian Li. A generalized CMOS-MEMS platform for micromechanical resonators monolithically integrated with

circuits. *Journal of Micromechanics and Microengineering*, 21(6):065012, 2011. URL: `http://stacks.iop.org/0960-1317/21/i=6/a=065012`.

[31] W. C. Chen, W. Fang, and S. S. Li. VHF CMOS-MEMS oxide resonators with Q > 10,000. In *Frequency Control Symposium (FCS), 2012 IEEE International*, pages 1–4, May 2012. `doi:10.1109/FCS.2012.6243706`.

[32] D. Weinstein and S. A. Bhave. Internal Dielectric Transduction in Bulk-Mode Resonators. *Journal of Microelectromechanical Systems*, 18(6):1401–1408, Dec 2009. `doi:10.1109/JMEMS.2009.2032480`.

[33] W. E. Newell. Face-mounted piezoelectric resonators. *Proceedings of the IEEE*, 53(6):575–581, June 1965. `doi:10.1109/PROC.1965.3925`.

[34] Wentao Wang and D. Weinstein. Acoustic Bragg reflectors for Q-enhancement of unreleased MEMS resonators. In *Frequency Control and the European Frequency and Time Forum (FCS), 2011 Joint Conference of the IEEE International*, pages 1–6, 2011. `doi:10.1109/FCS.2011.5977903`.

[35] Dana Weinstein and Sunil A. Bhave. The Resonant Body Transistor. *Nano Letters*, 10(4):1234–1237, 2010.

[36] R. Marathe, B. Bahr, W. Wang, Z. Mahmood, L. Daniel, and D. Weinstein. Resonant Body Transistors in IBM's 32 nm SOI CMOS Technology. *J. Microelectromech. Syst.*, (99):1–1, 2013. `doi:10.1109/JMEMS.2013.2283720`.

[37] A.M. Ionescu. Resonant body transistors. In *Device Research Conference (DRC), 2010*, pages 181–182, June 2010. `doi:10.1109/DRC.2010.5551901`.

[38] S.T. Bartsch, C. Dupre, E. Ollier, and A.M. Ionescu. Resonant-body silicon nanowire field effect transistor without junctions. In *Electron Devices Meeting (IEDM), 2012 IEEE International*, pages 15.2.1–15.2.4, Dec 2012. `doi:10.1109/IEDM.2012.6479046`.

[39] L.C. Popa and D. Weinstein. Switchable piezoelectric transduction in AlGaN/GaN MEMS resonators. In *Solid-State Sensors, Actuators and Microsystems (TRANSDUCERS EUROSENSORS XXVII), 2013 Transducers Eurosensors XXVII: The 17th International Conference on*, pages 2461–2464, June 2013. `doi:10.1109/Transducers.2013.6627304`.

[40] A. Ansari and M. Rais-Zadeh. HEMT-based read-out of a thickness-mode AlGaN/GaN resonator. In *Electron Devices Meeting (IEDM), 2013 IEEE International*, pages 18.3.1–18.3.4, Dec 2013. `doi:10.1109/IEDM.2013.6724653`.

[41] M. Faucher, Y. Cordier, M. Werquin, L. Buchaillot, C. Gaquiere, and D. Theron. Electromechanical Transconductance Properties of a GaN MEMS Resonator With Fully Integrated HEMT Transducers. *Microelectromechanical Systems, Journal of*, 21(2):370–378, April 2012. `doi:10.1109/JMEMS.2011.2179010`.

[42] F. NiebelschÃijtz, V. Cimalla, K. Tonisch, Ch. Haupt, K. BrÃijckner, R. Stephan, M. Hein, and O. Ambacher. AlGaN/GaN-based MEMS with two-dimensional electron gas for novel sensor applications. *physica status solidi (c)*, 5(6):1914–1916, 2008. doi:10.1002/pssc.200778424.

[43] M. Rais-Zadeh, V.J. Gokhale, A. Ansari, M. Faucher, D. TheÌĄron, Y. Cordier, and L. Buchaillot. Gallium Nitride as an Electromechanical Material. *Microelectromechanical Systems, Journal of*, 23(6):1252–1271, Dec 2014. doi:10.1109/JMEMS.2014.2352617.

[44] L. C. Popa. *Gallium Nitride MEMS Resonators*. PhD thesis, Massachusetts Institute of Technology, 2015.

[45] M. Faucher, Y. Cordier, M. Werquin, L. Buchaillot, C. Gaquiere, and D. Theron. Electromechanical Transconductance Properties of a GaN MEMS Resonator With Fully Integrated HEMT Transducers. *Microelectromechanical Systems, Journal of*, 21(2):370–378, April 2012. doi:10.1109/JMEMS.2011.2179010.

[46] Lei Shao, Meng Zhang, Animesh Banerjee, Pallab Bhattacharya, and Kevin P. Pipe. Emission and detection of surface acoustic waves by AlGaN/GaN high electron mobility transistors. *Applied Physics Letters*, 99(24), 2011. URL: http://scitation.aip.org/content/aip/journal/apl/99/24/10.1063/1.3665625, doi:http://dx.doi.org/10.1063/1.3665625.

[47] L.C. Popa and D. Weinstein. Switchable piezoelectric transduction in AlGaN/GaN MEMS resonators. In *Solid-State Sensors, Actuators and Microsystems (TRANSDUCERS EUROSENSORS XXVII), 2013 Transducers Eurosensors XXVII: The 17th International Conference on*, pages 2461–2464, June 2013. doi:10.1109/Transducers.2013.6627304.

[48] L.C. Popa and D. Weinstein. 2DEG electrodes for piezoelectric transduction of AlGaN/GaN MEMS resonators. In *European Frequency and Time Forum International Frequency Control Symposium (EFTF/IFC), 2013 Joint*, pages 922–925, July 2013. doi:10.1109/EFTF-IFC.2013.6702242.

[49] L.C. Popa and D. Weinstein. 1 GHz GaN resonant body transistors with enhanced off-resonance rejection. *Solid-State Sensors, Actuators, and Microsystems Workshop (Hilton Head 2014)*, pages 269–272, 2014.

[50] A. Ansari and M. Rais-Zadeh. An 8.7 GHz GaN Micromechnaical Resonator With an Integrated AlGaN/GaN HEMT. In *Solid-State Sensors, Actuators and Microsystems Workshop*, 2014.

[51] A. Ansari and M. Rais-Zadeh. A Temperature-Compensated Gallium Nitride Micromechanical Resonator. *Electron Device Letters, IEEE*, 35(11):1127–1129, Nov 2014. doi:10.1109/LED.2014.2358577.

[52] B.W. Bahr, L.C. Popa, and D. Weinstein. 1GHz GaN-MMIC monolithically integrated MEMS-based oscillators. In *Solid- State Circuits Conference - (ISSCC), 2015 IEEE International*, pages 1–3, Feb 2015. doi:10.1109/ISSCC.2015.7063047.

[53] M. S. Kushwaha, P. Halevi, L. Dobrzynski, and B. Djafari-Rouhani. Acoustic band structure of periodic elastic composites. *Phys. Rev. Lett.*, 71:2022–2025, 1993.

[54] M. S. Kushwaha, P. Halevi, G. Martinez, L. Dobrzynski, and B. Djafari-Rouhani. Theory of acoustic band structure of periodic elastic composites. *Phys. Rev. B*, 49:2313–2322, Jan 1994. doi:10.1103/PhysRevB.49.2313.

[55] I. El-Kady, R. H. Olsson, and J. G. Fleming. Phononic band-gap crystals for radio frequency communications. *Applied Physics Letters*, 92(23), 2008. URL: http://scitation.aip.org/content/aip/journal/apl/92/23/10.1063/1.2938863, doi:http://dx.doi.org/10.1063/1.2938863.

[56] Charles M. Reinke, M. F. Su, R. H. Olsson, and I. El-Kady. Realization of optimal bandgaps in solid-solid, solid-air, and hybrid solid-air-solid phononic crystal slabs. *Applied Physics Letters*, 98(6), 2011. URL: http://scitation.aip.org/content/aip/journal/apl/98/6/10.1063/1.3543848, doi:http://dx.doi.org/10.1063/1.3543848.

[57] M. F. Su, R. H. Olsson, Z. C. Leseman, and I. El-Kady. Realization of a phononic crystal operating at gigahertz frequencies. *Applied Physics Letters*, 96(5), 2010. URL: http://scitation.aip.org/content/aip/journal/apl/96/5/10.1063/1.3280376, doi:http://dx.doi.org/10.1063/1.3280376.

[58] Yan-Feng Wang, Yue-Sheng Wang, and Xiao-Xing Su. Large bandgaps of two-dimensional phononic crystals with cross-like holes. *Journal of Applied Physics*, 110(11):113520, 2011. doi:10.1063/1.3665205.

[59] B. Bahr, R. Marathe, and D. Weinstein. Phononic crystals for acoustic confinement in CMOS-MEMS resonators. In *Frequency Control Symposium (FCS), 2014 IEEE International*, pages 1–4, May 2014. doi:10.1109/FCS.2014.6859980.

[60] M. Ziaei-Moayyed, M.F. Su, C.M. Reinke, El-Kady, and R.H. Olsson. Silicon carbide phononic crystals for high f.Q micromechanical resonators. In *Ultrasonics Symposium (IUS), 2010 IEEE*, pages 162–166, Oct 2010. doi:10.1109/ULTSYM.2010.5935682.

[61] Nai-Kuei Kuo and G. Piazza. 1 GHZ phononic band gap structure in air/aluminum nitride for symmetric lamb waves. In *Micro Electro Mechanical Systems (MEMS), 2011 IEEE 24th International Conference on*, pages 740–743, Jan 2011. doi:10.1109/MEMSYS.2011.5734531.

[62] S. Wang, L.C. Popa, and D. Weinstein. GaN MEMS resonator using a folded phononic crystal structure. In *Solid-State Sensors, Actuators, and Microsystems Workshop (Hilton Head 2014)*, pages 72–75, 2014.

[63] Cheong Yang Koh. *Generalized phononic networks : of length scales, symmetry breaking and (non) locality.* PhD thesis, Massachusetts Institute of Technology, 2011.

[64] Tsung-Tsong Wu, Yan-Ting Chen, Jia-Hong Sun, Sz-Chin Steven Lin, and Tony Jun Huang. Focusing of the lowest antisymmetric Lamb wave in a gradient-index phononic crystal plate. *Applied Physics Letters*, 98(17), 2011. URL: `http://scitation.aip.org/content/aip/journal/apl/98/17/10.1063/1.3583660`, `doi:http://dx.doi.org/10.1063/1.3583660`.

[65] R. H. Olsson, M. Ziaei-Moayyed, B. Kim, C. Reinke, M. F. Su, P. Hopkins, Y. M. Soliman, D. F. Goettler, Z. C. Leseman, and I. El-Kady. Micro and nano fabricated phononic crystals: technology and applications. In *Ultrasonics Symposium (IUS), 2011 IEEE International*, pages 983–988, Oct 2011. `doi:10.1109/ULTSYM.2011.0241`.

[66] A. Khelif, A. Choujaa, B. Djafari-Rouhani, M. Wilm, S. Ballandras, and V. Laude. Trapping and guiding of acoustic waves by defect modes in a full-band-gap ultrasonic crystal. *Phys. Rev. B*, 68:214301, Dec 2003. URL: `http://link.aps.org/doi/10.1103/PhysRevB.68.214301`, `doi:10.1103/PhysRevB.68.214301`.

[67] Saeed Mohammadi, Ali Asghar Eftekhar, William D. Hunt, and Ali Adibi. High-Q micromechanical resonators in a two-dimensional phononic crystal slab. *Applied Physics Letters*, 94(5), 2009. URL: `http://scitation.aip.org/content/aip/journal/apl/94/5/10.1063/1.3078284`, `doi:http://dx.doi.org/10.1063/1.3078284`.

[68] B.A. Auld. *Acoustic Fields and Waves in Solids: vol. 1.* Acoustic Fields and Waves in Solids. Krieger Publishing Company, 1990.

[69] J.J. Sakurai and J. Napolitano. *Modern Quantum Mechanics.* Addison-Wesley, 2011. URL: `https://books.google.com/books?id=N4I-AQAACAAJ`.

[70] Siping Wang, Wen-Chien Chen, Bichoy Bahr, Weileun Fang, Sheng-Shian Li, and Dana Weinstein. Temperature coefficient of frequency modeling for CMOS-MEMS bulk mode composite resonators. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 62(6):1166–1178, June 2015. `doi:10.1109/TUFFC.2014.006724`.

[71] J.D. Joannopoulos, S.G. Johnson, J.N. Winn, and R.D. Meade. *Photonic Crystals: Molding the Flow of Light (Second Edition).* Princeton University Press, 2011.

[72] P. Mendelson J. S. Lomont. The Wigner Unitarity-Antiunitarity Theorem. *Annals of Mathematics*, 78(3):548–559, 1963. URL: `http://www.jstor.org/stable/1970540`.

[73] J.L. Safko, H. Goldstein, and C.P. Poole. *Classical Mechanics*. Pearson Education, 2002. URL: `https://books.google.com/books?id=Spy6xHWFJIEC`.

[74] C. Kittel. *Introduction to Solid State Physics*. Wiley, 2004.

[75] B. Greene et al. High performance 32nm SOI CMOS with high-k/metal gate and 0.149μm^2 SRAM and ultra low-k back end with eleven levels of copper. In *VLSI Technology, 2009 Symposium on*, pages 140–141, 2009.

[76] COMSOL. *version 4.4*. COMSOL, Inc., 2014. URL: `http://www.comsol.com`.

[77] Jean-Pierre Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics*, 114(2):185 – 200, 1994. URL: `http://www.sciencedirect.com/science/article/pii/S0021999184711594`, `doi:http://dx.doi.org/10.1006/jcph.1994.1159`.

[78] Steven G. Johnson, M. Ibanescu, M. A. Skorobogatiy, O. Weisberg, J. D. Joannopoulos, and Y. Fink. Perturbation theory for Maxwell's equations with shifting material boundaries. *Phys. Rev. E*, 65:066611, Jun 2002. URL: `http://link.aps.org/doi/10.1103/PhysRevE.65.066611`, `doi:10.1103/PhysRevE.65.066611`.

[79] B. Bahr, R. Marathe, and D. Weinstein. Theory and Design of Phononic Crystals for Unreleased CMOS-MEMS Resonant Body Transistors. *Microelectromechanical Systems, Journal of*, PP(99):1–1, 2015. `doi:10.1109/JMEMS.2015.2418789`.

[80] B. Bahr and D. Weinstein. Vertical Acoustic Confinement for High-Q Fully-Differential CMOS-RBTs. In *Solid-State Sensors, Actuators, and Microsystems Workshop (Hilton Head 2016)*, page p.p., 2016.

[81] B. Gustavsen and A. Semlyen. Rational approximation of frequency domain responses by vector fitting. *Power Delivery, IEEE Transactions on*, 14(3):1052–1061, Jul 1999. `doi:10.1109/61.772353`.

[82] R. Tabrizian, M. Rais-Zadeh, and F. Ayazi. Effect of phonon interactions on limiting the f.Q product of micromechanical resonators. In *Solid-State Sensors, Actuators and Microsystems Conference, 2009. TRANSDUCERS 2009. International*, pages 2131–2134, June 2009. `doi:10.1109/SENSOR.2009.5285627`.

[83] Jian-Ming Jin. *Theory and Computation of Electromagnetic Fields*. Wiley-Blackwell, sep 2010. URL: `http://dx.doi.org/10.1002/9780470874257`, `doi:10.1002/9780470874257`.

[84] Alfred Grill. Porous pSiCOH Ultralow-k Dielectrics for Chip Interconnects Prepared by PECVD. *Annual Review of Materials Research*, 39(1):49–69, 2009. doi:10.1146/annurev-matsci-082908-145305.

[85] G. Degrande, G.De Roeck, P.Van Den Broeck, and D. Smeulders. Wave propagation in layered dry, saturated and unsaturated poroelastic media. *International Journal of Solids and Structures*, 35(34-35):4753 – 4778, 1998. doi:http://dx.doi.org/10.1016/S0020-7683(98)00093-6.

[86] Steven G. Johnson, Peter Bienstman, M. A. Skorobogatiy, Mihai Ibanescu, Elefterios Lidorikis, and J. D. Joannopoulos. Adiabatic theorem and continuous coupled-mode theory for efficient taper transitions in photonic crystals. *Phys. Rev. E*, 66:066608, Dec 2002. doi:10.1103/PhysRevE.66.066608.

[87] S.G. Johnson, M.L. Povinelli, P. Bienstman, M. Skorobogatiy, M. Soljacic, M. Ibanescu, E. Lidorikis, and J.D. Joannopoulos. Coupling, scattering, and perturbation theory: Semi-analytical analyses of photonic-crystal waveguides. In *Transparent Optical Networks, 2003. Proceedings of 2003 5th International Conference on*, volume 1, pages 103–109 vol.1, June 2003. doi:10.1109/ICTON.2003.1264589.

[88] Jeroen Tromp. A Coupled Local-Mode Analysis of Surface-Wave Propagation In A Laterally Heterogeneous Waveguide. *Geophysical Journal International*, 117(1):153–161, 1994. doi:10.1111/j.1365-246X.1994.tb03309.x.

[89] B.E. Little and H.A. Haus. A variational coupled-mode theory for periodic waveguides. *Quantum Electronics, IEEE Journal of*, 31(12):2258–2264, Dec 1995. doi:10.1109/3.477755.

[90] L.D. Landau and E.M. Lifshitïÿăsïÿą. *Quantum Mechanics: Non-relativistic Theory*. Butterworth Heinemann. Butterworth-Heinemann, 1977. URL: https://books.google.com/books?id=J9ui6KwC4mMC.

[91] Albert Messiah. *Quantum Mechanics*. Number v. 2 in Quantum Mechanics. North-Holland, 1965. URL: https://books.google.com/books?id=VR93vUk8d_8C.

[92] E. Joseph Avron and Alexander Elgart. Adiabatic Theorem without a Gap Condition. *Communications in Mathematical Physics*, 203(2):445–463, 1999. URL: http://dx.doi.org/10.1007/s002200050620, doi:10.1007/s002200050620.

[93] J. K. Chandalia, B. J. Eggleton, R. S. Windeler, S. G. Kosinski, X. Liu, and C. Xu. Adiabatic coupling in tapered air-silica microstructured optical fiber. *IEEE Photonics Technology Letters*, 13(1):52–54, Jan 2001. doi:10.1109/68.903218.

[94] W. Bogaerts and S. K. Selvaraja. Compact Single-Mode Silicon Hybrid Rib/Strip Waveguide With Adiabatic Bends. *IEEE Photonics Journal*, 3(3):422–432, June 2011. doi:10.1109/JPHOT.2011.2142931.

[95] K. Ustun and H. Kurt. Efficient and compact coupling to slow light structures. In *Optical MEMS and Nanophotonics (OMN), 2011 International Conference on*, pages 105–106, Aug 2011. doi:10.1109/OMEMS.2011.6031091.

[96] C. Y. Lin, Z. Xue, A. X. Wang, and R. T. Chen. Wideband and group index independent coupling to slow light slotted photonic crystal waveguides with adiabatic group index taper and mode matching. In *Lasers and Electro-Optics (CLEO), 2012 Conference on*, pages 1–2, May 2012.

[97] Ardavan Oskooi, Almir Mutapcic, Susumu Noda, J. D. Joannopoulos, Stephen P. Boyd, and Steven G. Johnson. Robust optimization of adiabatic tapers for coupling to slow-light photonic-crystal waveguides. *Opt. Express*, 20(19):21558–21575, Sep 2012. URL: http://www.opticsexpress.org/abstract.cfm?URI=oe-20-19-21558, doi:10.1364/OE.20.021558.

[98] S. Wang, L. C. Popa, and D. Weinstein. Tapered Phononic Crystal sawresonator in GaN. In *Micro Electro Mechanical Systems (MEMS), 2015 28th IEEE International Conference on*, pages 1028–1031, Jan 2015. doi:10.1109/MEMSYS.2015.7051137.

[99] Guang-Hua Duan, C. Jany, A. Le liepvre, A. Accard, M. Lamponi, D. Make, P. Kaspar, G. Levaufre, N. Girard, F. Lelarge, J. M. Fedeli, A. Descos, B. Ben Bakir, S. Messaoudene, D. Bordel, S. Menezo, G. de Valicourt, S. Keyvaninia, G. Roelkens, D. Van Thourhout, D. J. Thomson, F. Y. Gardes, and G. T. Reed. Hybrid III–V on Silicon Lasers for Photonic Integrated Circuits on Silicon. *IEEE Journal of Selected Topics in Quantum Electronics*, 20(4):158–170, July 2014. doi:10.1109/JSTQE.2013.2296752.

[100] Hannah Clevenson, Pierre Desjardins, Xuetao Gan, and Dirk Englund. High sensitivity gas sensor based on high-Q suspended polymer photonic crystal nanocavity. *Applied Physics Letters*, 104(24), 2014. URL: http://scitation.aip.org/content/aip/journal/apl/104/24/10.1063/1.4879735, doi:http://dx.doi.org/10.1063/1.4879735.

[101] MATLAB. *version R2014a*. The MathWorks Inc., Natick, Massachusetts, 2014.

[102] H. Richard Byrd, Charles Jean Gilbert, and Jorge Nocedal. A trust region method based on interior point techniques for nonlinear programming. *Mathematical Programming*, 89(1):149–185, 2000. URL: http://dx.doi.org/10.1007/PL00011391, doi:10.1007/PL00011391.

[103] L.T. Pillage and R.A. Rohrer. Asymptotic waveform evaluation for timing analysis. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 9(4):352–366, Apr 1990. doi:10.1109/43.45867.

[104] M. S. Lundstrom. On the mobility versus drain current relation for a nanoscale MOSFET. *IEEE Electron Device Letters*, 22(6):293–295, June 2001. doi:10.1109/55.924846.

[105] Cascade Microtech® - https://www.cascademicrotech.com/.

[106] E. J. Nowak, I. Aller, T. Ludwig, K. Kim, R. V. Joshi, Ching-Te Chuang, K. Bernstein, and R. Puri. Turning silicon on its edge [double gate CMOS/FinFET technology]. *IEEE Circuits and Devices Magazine*, 20(1):20–31, Jan 2004. doi:10.1109/MCD.2004.1263404.

[107] Bichoy Bahr, Dana Weinstein, and Luca Daniel. Released Resonant Body Transistor (RBT) Model, Oct 2014. URL: https://nanohub.org/publications/17/1, doi:doi:/10.4231/D3KS6J55W.

[108] Bichoy Bahr, Dana Weinstein, and Luca Daniel. Released Resonant Body Transistor with MIT Virtual Source (RBT-MVS) Model, Aug 2015. URL: https://nanohub.org/publications/72/1, doi:doi:/10.4231/D3VH5CK04.

[109] Bichoy Bahr, Dana Weinstein, and Luca Daniel. Unreleased 1D CMOS Resonant Body Transistor
with MIT Virtual Source (URBT-MVS) Compact Model, March 2016. URL: https://nanohub.org/publications/132/1.

[110] BSIM models. URL: http://www-device.eecs.berkeley.edu/bsim/.

[111] Shaloo Rakheja and Dimitri Antoniadis. MVS Nanotransistor Model (Silicon), Aug 2015. URL: https://nanohub.org/publications/15, doi:doi:/10.4231/D3QZ22J40.

[112] C. Zuo, N. Sinha, J. Van der Spiegel, and G. Piazza. Multifrequency Pierce Oscillators Based on Piezoelectric AlN Contour-Mode MEMS Technology. *Journal of Microelectromechanical Systems*, 19(3):570–580, June 2010. doi:10.1109/JMEMS.2010.2045879.

[113] M. Rinaldi, C. Zuo, J. Van der Spiegel, and G. Piazza. Reconfigurable CMOS Oscillator Based on Multifrequency AlN Contour-Mode MEMS Resonators. *IEEE Transactions on Electron Devices*, 58(5):1281–1286, May 2011. doi:10.1109/TED.2011.2104961.

[114] Jabeom Koo, A. Tazzoli, J. Segovai-Fernandez, G. Piazza, and B. Otis. A -173dBc/Hz @ 1 MHz offset Colpitts oscillator using AlN contour-mode MEMS resonator. In *Custom Integrated Circuits Conference (CICC), 2013 IEEE*, pages 1–4, Sept 2013. doi:10.1109/CICC.2013.6658547.

[115] H.M. Lavasani, Wanling Pan, B. Harrington, R. Abdolvand, and F. Ayazi. A 76 dBΩ; 1.7GHz 0.18 μm CMOS tunable transimpedance amplifier using broadband current pre-amplifier for high frequency lateral micromechanical oscillators. In *Solid-State Circuits Conference Digest of Technical Papers (ISSCC),*

2010 IEEE International, pages 318–319, Feb 2010. doi:10.1109/ISSCC.2010.5433902.

[116] S. Razafimandimby, A. Cathelin, J. Lajoinie, A. Kaiser, and D. Belot. A 2GHz 0.25μm SiGe BiCMOS Oscillator with Flip-Chip Mounted BAW Resonator. In *Solid-State Circuits Conference, 2007. ISSCC 2007. Digest of Technical Papers. IEEE International*, pages 580–623, Feb 2007. doi:10.1109/ISSCC.2007.373553.

[117] M. Aissi, E. Tournier, M.-A. Dubois, G. Parat, and R. Plana. A 5.4GHz 0.35 μm BiCMOS FBAR Resonator Oscillator in Above-IC Technology. In *Solid-State Circuits Conference, 2006. ISSCC 2006. Digest of Technical Papers. IEEE International*, pages 1228–1235, Feb 2006. doi:10.1109/ISSCC.2006.1696169.

[118] O. Ambacher, J. Smart, J. R. Shealy, N. G. Weimann, K. Chu, M. Murphy, W. J. Schaff, L. F. Eastman, R. Dimitrov, L. Wittmer, M. Stutzmann, W. Rieger, and J. Hilsenbeck. Two-dimensional electron gases induced by spontaneous and piezoelectric polarization charges in N- and Ga-face AlGaN/GaN heterostructures. *Journal of Applied Physics*, 85(6):3222–3233, 1999. URL: http://scitation.aip.org/content/aip/journal/jap/85/6/10.1063/1.369664, doi:http://dx.doi.org/10.1063/1.369664.

[119] R. Gaska, M. S. Shur, A. D. Bykhovski, A. O. Orlov, and G. L. Snider. Electron mobility in modulation-doped AlGaNâĂŞGaN heterostructures. *Applied Physics Letters*, 74(2):287–289, 1999. URL: http://scitation.aip.org/content/aip/journal/apl/74/2/10.1063/1.123001, doi:http://dx.doi.org/10.1063/1.123001.

[120] F. Di Pietrantonio, M. Benetti, D. Cannata, R. Beccherelli, and E. Verona. Guided lamb wave electroacoustic devices on micromachined AlN/Al plates. *Ultrasonics, Ferroelectrics, and Frequency Control, IEEE Transactions on*, 57(5):1175–1182, May 2010. doi:10.1109/TUFFC.2010.1530.

[121] Chih-Ming Lin, V. Yantchev, Jie Zou, Yung-Yu Chen, and A.P. Pisano. Micromachined One-Port Aluminum Nitride Lamb Wave Resonators Utilizing the Lowest-Order Symmetric Mode. *Microelectromechanical Systems, Journal of*, 23(1):78–91, Feb 2014. doi:10.1109/JMEMS.2013.2290793.

[122] Qiang Chen. Latest advances in gallium nitride HEMT modeling. In *Solid-State and Integrated Circuit Technology (ICSICT), 2014 12th IEEE International Conference on*, pages 1–4, Oct 2014. doi:10.1109/ICSICT.2014.7021208.

[123] L. Dunleavy, C. Baylis, W. Curtice, and R. Connick. Modeling GaN: Powerful but Challenging. *Microwave Magazine, IEEE*, 11(6):82–96, Oct 2010. doi:10.1109/MMM.2010.937735.

[124] I. Angelov, H. Zirath, and N. Rosman. A new empirical nonlinear model for HEMT and MESFET devices. *Microwave Theory and Techniques, IEEE Transactions on*, 40(12):2258–2266, Dec 1992. `doi:10.1109/22.179888`.

[125] S. Khandelwal, C. Yadav, S. Agnihotri, Y.S. Chauhan, A. Curutchet, T. Zimmer, J.-C. De Jaeger, N. Defrance, and T.A. Fjeldly. Robust Surface-Potential-Based Compact Model for GaN HEMT IC Design. *Electron Devices, IEEE Transactions on*, 60(10):3216–3222, Oct 2013. `doi:10.1109/TED.2013.2265320`.

[126] P. Martin and L. Lucci. A compact model of AlGaN/GaN HEMTs power transistors based on a surface-potential approach. In *Mixed Design of Integrated Circuits and Systems (MIXDES), 2013 Proceedings of the 20th International Conference*, pages 92–95, June 2013.

[127] P. Martin and L. Lucci. A compact model of AlGaN/GaN HEMTs power transistors based on a surface-potential approach. In *Mixed Design of Integrated Circuits and Systems (MIXDES), 2013 Proceedings of the 20th International Conference*, pages 92–95, June 2013.

[128] U. Radhakrishna, D. Piedra, Yuhao Zhang, T. Palacios, and D. Antoniadis. High voltage GaN HEMT compact model: Experimental verification, field plate optimization and charge trapping. In *Electron Devices Meeting (IEDM), 2013 IEEE International*, pages 32.7.1–32.7.4, Dec 2013. `doi:10.1109/IEDM.2013.6724740`.

[129] Ujwal Radhakrishna. *A compact transport and charge model for GaN-based high electron mobility transistors for RF applications.* PhD thesis, Massachusetts Institute of Technology, 2013.

[130] Ujwal Radhakrishna and Dimitri Antoniadis. MIT Virtual Source GaNFET-RF (MVSG-RF) Model, Oct 2014. URL: `https://nanohub.org/publications/14`, `doi:doi:/10.4231/D3G15TC12`.

[131] Jianjun Xu, J. Horn, M. Iwamoto, and D.E. Root. Large-signal FET model with multiple time scale dynamics from nonlinear vector network analyzer data. In *Microwave Symposium Digest (MTT), 2010 IEEE MTT-S International*, pages 417–420, May 2010. `doi:10.1109/MWSYM.2010.5516843`.

[132] W.R. Curtice and M. Ettenberg. A Nonlinear GaAs FET Model for Use in the Design of Output Circuits for Power Amplifiers. In *Microwave Symposium Digest, 1985 IEEE MTT-S International*, pages 405–408, June 1985. `doi:10.1109/MWSYM.1985.1131996`.

[133] Valiallah Zomorrodian, Yi Pei, Umesh K. Mishra, and Robert A. York. A scalable EE-HEMT based large signal model for multi-finger AlGaN/GaN HEMTs for linear and non-linear circuit design. *physica status solidi (c)*,

7(10):2450–2454, 2010. URL: http://dx.doi.org/10.1002/pssc.200983877, doi:10.1002/pssc.200983877.

[134] G. Meneghesso, G. Verzellesi, Roberto Pierobon, F. Rampazzo, A. Chini, Umesh K. Mishra, C. Canali, and E. Zanoni. Surface-related drain current dispersion effects in AlGaN-GaN HEMTs. *Electron Devices, IEEE Transactions on*, 51(10):1554–1561, Oct 2004. doi:10.1109/TED.2004.835025.

[135] R. Vetury, N.Q. Zhang, Stacia Keller, and Umesh K. Mishra. The impact of surface states on the DC and RF characteristics of AlGaN/GaN HFETs. *Electron Devices, IEEE Transactions on*, 48(3):560–566, Mar 2001. doi:10.1109/16.906451.

[136] G. Meneghesso, F. Rampazzo, P. Kordos, G. Verzellesi, and E. Zanoni. Current Collapse and High-Electric-Field Reliability of Unpassivated GaN/AlGaN/GaN HEMTs. *Electron Devices, IEEE Transactions on*, 53(12):2932–2941, Dec 2006. doi:10.1109/TED.2006.885681.

[137] A. Jarndal and G. Kompa. Large-Signal Model for AlGaN/GaN HEMTs Accurately Predicts Trapping- and Self-Heating-Induced Dispersion and Intermodulation Distortion. *Electron Devices, IEEE Transactions on*, 54(11):2830–2836, Nov 2007. doi:10.1109/TED.2007.907143.

[138] J.M. Tirado, J.L. Sanchez-Rojas, and J.I. Izpura. Trapping Effects in the Transient Response of AlGaN/GaN HEMT Devices. *Electron Devices, IEEE Transactions on*, 54(3):410–417, March 2007. doi:10.1109/TED.2006.890592.

[139] Feng Gao, Di Chen, Bin Lu, H.L. Tuller, C.V. Thompson, S. Keller, U.K. Mishra, and T. Palacios. Impact of Moisture and Fluorocarbon Passivation on the Current Collapse of AlGaN/GaN HEMTs. *Electron Device Letters, IEEE*, 33(10):1378–1380, Oct 2012. doi:10.1109/LED.2012.2206556.

[140] H. F. Sun and C. R. Bolognesi. Anomalous behavior of AlGaNâĹŢGaN heterostructure field-effect transistors at cryogenic temperatures: From current collapse to current enhancement with cooling. *Applied Physics Letters*, 90(12), 2007. URL: http://scitation.aip.org/content/aip/journal/apl/90/12/10.1063/1.2715032, doi:http://dx.doi.org/10.1063/1.2715032.

[141] G. Meneghesso, F. Zanon, M.J. Uren, and E. Zanoni. Anomalous Kink Effect in GaN High Electron Mobility Transistors. *Electron Device Letters, IEEE*, 30(2):100–102, Feb 2009. doi:10.1109/LED.2008.2010067.

[142] G. Meneghesso, F. Rossi, G. Salviati, M. J. Uren, E. MuÃśoz, and E. Zanoni. Correlation between kink and cathodoluminescence spectra in AlGaN/GaN high electron mobility transistors. *Applied Physics Letters*, 96(26), 2010. URL: http://scitation.aip.org/content/aip/journal/apl/96/26/10.1063/1.3459968, doi:http://dx.doi.org/10.1063/1.3459968.

[143] Maojun Wang and K.J. Chen. Kink Effect in AlGaN/GaN HEMTs Induced by Drain and Gate Pumping. *Electron Device Letters, IEEE*, 32(4):482–484, April 2011. doi:10.1109/LED.2011.2105460.

[144] Behzad Razavi. *Design of Analog CMOS Integrated Circuits*. McGraw-Hill, Inc., New York, NY, USA, 1 edition, 2001.

[145] E.A. Vittoz, M.G.R. Degrauwe, and S. Bitz. High-performance crystal oscillator circuits: theory and application. *Solid-State Circuits, IEEE Journal of*, 23(3):774–783, June 1988. doi:10.1109/4.318.

[146] IEEE. IEEE Standard on Piezoelectricity. *ANSI/IEEE Std 176-1987*, 1988. doi:10.1109/IEEESTD.1988.79638.

[147] III. Larson, J.D., R.C. Bradley, S. Wartenberg, and R.C. Ruby. Modified Butterworth-Van Dyke circuit for FBAR resonators and automated measurement system. In *Ultrasonics Symposium, 2000 IEEE*, volume 1, pages 863–868 vol.1, Oct 2000. doi:10.1109/ULTSYM.2000.922679.

[148] Mathematica® - Wolfram https://www.wolfram.com/.

[149] F. Silveira, D. Flandre, and P. G. A. Jespers. A gm/ID based methodology for the design of CMOS analog circuits and its application to the synthesis of a silicon-on-insulator micropower OTA. *IEEE Journal of Solid-State Circuits*, 31(9):1314–1319, Sep 1996. doi:10.1109/4.535416.

[150] S. Giraud, S. Bila, M. Chatras, D. Cros, and M. Aubourg. Bulk acoustic wave filter synthesis and optimization for UMTS applications. In *Microwave Conference, 2009. EuMC 2009. European*, pages 456–459, Sept 2009. doi:10.1109/FREQ.2009.5168315.

[151] R. Lanz and P. Muralt. Solidly mounted BAW filters for 8 GHz based on AlN thin films. In *Ultrasonics, 2003 IEEE Symposium on*, volume 1, pages 178–181 Vol.1, Oct 2003. doi:10.1109/ULTSYM.2003.1293383.

[152] B. Kim, R. H. Olsson, and K. E. Wojciechowski. AlN Microresonator-Based Filters With Multiple Bandwidths at Low Intermediate Frequencies. *Journal of Microelectromechanical Systems*, 22(4):949–961, Aug 2013. doi:10.1109/JMEMS.2013.2251414.

[153] D. Hatanaka, I. Mahboob, K. Onomitsu, and H. Yamaguchi. Phonon waveguides for electromechanical circuits. *Nat Nano*, 9(7):520–524, Jul 2014. Letter. URL: http://dx.doi.org/10.1038/nnano.2014.107.

[154] S. Summerfelt, S. Aggarwal, K. Boku, F. Celii, L. Hall, L. Matz, S. Martin, H. McAdams, K. Remack, J. Rodriguez, K. Taylor, K. R. Udayakumar, T. Moise, R. Bailey, M. Depner, G. Fox, and J. Eliason. Embedded ferroelectric memory using a 130-nm 5 metal layer Cu / FSG logic process. In

Non-Volatile Memory Technology Symposium, 2004, pages 153–154, Nov 2004. doi:10.1109/NVMT.2004.1380833.

[155] J. MÃijller, E. Yurchuk, T. SchlÃűsser, J. Paul, R. Hoffmann, S. MÃijller, D. Martin, S. Slesazeck, P. Polakowski, J. Sundqvist, M. Czernohorsky, K. Seidel, P. KÃijcher, R. Boschke, M. Trentzsch, K. Gebauer, U. SchrÃűder, and T. Mikolajick. Ferroelectricity in HfO2 enables nonvolatile data storage in 28 nm HKMG. In *VLSI Technology (VLSIT), 2012 Symposium on*, pages 25–26, June 2012. doi:10.1109/VLSIT.2012.6242443.

[156] J. MÃijller, T. S. BÃűscke, S. MÃijller, E. Yurchuk, P. Polakowski, J. Paul, D. Martin, T. Schenk, K. Khullar, A. Kersch, W. Weinreich, S. Riedel, K. Seidel, A. Kumar, T. M. Arruda, S. V. Kalinin, T. SchlÃűsser, R. Boschke, R. van Bentum, U. SchrÃűder, and T. Mikolajick. Ferroelectric hafnium oxide: A CMOS-compatible and highly scalable approach to future ferroelectric memories. In *Electron Devices Meeting (IEDM), 2013 IEEE International*, pages 10.8.1–10.8.4, Dec 2013. doi:10.1109/IEDM.2013.6724605.

[157] P. Maffezzoni, B. Bahr, Z. Zhang, and L. Daniel. Analysis and Design of Weakly Coupled LC Oscillator Arrays Based on Phase-Domain Macromodels. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 34(1):77–85, Jan 2015. doi:10.1109/TCAD.2014.2365360.

[158] P. Maffezzoni, B. Bahr, Z. Zhang, and L. Daniel. Oscillator Array Models for Associative Memory and Pattern Recognition. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 62(6):1591–1598, June 2015. doi:10.1109/TCSI.2015.2418851.

[159] P. Maffezzoni, B. Bahr, Z. Zhang, and L. Daniel. Reducing Phase Noise in Multi-Phase Oscillators. *IEEE Transactions on Circuits and Systems I: Regular Papers*, PP(99):1–10, 2016. doi:10.1109/TCSI.2016.2525078.

[160] Narayan Srinivasa et al. Probabilistic inference devices for unconventional processing of signals for intelligent data exploitation. In *GOMAC Tech 2015*, page 17.2, March 2015.

[161] Chen Sun et al. Single-chip microprocessor that communicates directly using light. *Nature*, 528(7583):534–538, Dec 2015. Letter. URL: http://dx.doi.org/10.1038/nature16454.

[162] Heedeuk Shin, Wenjun Qiu, Robert Jarecki, Jonathan A. Cox, Roy H. Olsson, Andrew Starbuck, Zheng Wang, and Peter T. Rakich. Tailorable stimulated Brillouin scattering in nanoscale silicon waveguides. *Nature Communications*, 4, jun 2013. URL: http://dx.doi.org/10.1038/ncomms2943, doi:10.1038/ncomms2943.

[163] Markus Aspelmeyer, Tobias J. Kippenberg, and Florian Marquardt. Cavity optomechanics. *Reviews of Modern Physics*, 86(4):1391–1452, dec

2014. URL: `http://dx.doi.org/10.1103/RevModPhys.86.1391`, `doi:10.1103/revmodphys.86.1391`.

[164] H. Baltes, O. Brand, G.K. Fedder, C. Hierold, J.G. Korvink, and O. Tabata. *CMOS-MEMS: Advanced Micro and Nanosystems*. Advanced Micro and Nanosystems. Wiley, 2008.