UNIVERSITY OF CALIFORNIA,
IRVINE

Mode Ordering in Anti-Phase Driven MEMS Gyroscopes and Accelerometers

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Mechanical and Aerospace Engineering

by

Brenton Ross Simon

Dissertation Committee:
Professor Andrei Shkel, Chair
Professor Faryar Jabbari
Dr. Alexander Trusov

2014
DEDICATION

To my family and friends whom I have ignored for far too long.
# Table of Contents

List of Figures ............................................................................................................. ix

List of Tables ............................................................................................................... xix

Acknowledgements ..................................................................................................... xx

Curriculum Vitae ......................................................................................................... xxi

Abstract of the Dissertation ....................................................................................... xxiv

1 Introduction ............................................................................................................... 1
  1.1 Motivation ............................................................................................................. 1
  1.2 Background ......................................................................................................... 3
    1.2.1 Principles of Operation .................................................................................. 3
    1.2.2 MEMS Gyroscope Materials ....................................................................... 5
    1.2.3 Types of Vibratory Gyroscopes ................................................................... 7
  1.3 Trends and Challenges ....................................................................................... 10
    1.3.1 Terminology ................................................................................................. 10
    1.3.2 Stability ......................................................................................................... 16
    1.3.3 Environmental Robustness .......................................................................... 18
  1.4 MEMS Inertial Sensor Prior Art ........................................................................ 20
    1.4.1 Commercial Devices .................................................................................... 20
    1.4.2 Academic Publications ................................................................................ 25
  1.5 Research Objective ............................................................................................. 31
  1.6 Dissertation Outline ........................................................................................... 32

2 Mode Ordering of Multi-Mass Resonators ............................................................... 35
  2.1 Introduction ......................................................................................................... 35
  2.2 Resonator Approach ........................................................................................... 36
    2.2.1 Distributed Mass ......................................................................................... 36
    2.2.2 Lumped Parameters .................................................................................... 38
  2.3 Motivation ............................................................................................................ 39
    2.3.1 Common-mode Rejection ............................................................................. 40
    2.3.2 Energy Loss through Structural Asymmetry .............................................. 43
    2.3.3 Linear Range Expansion ............................................................................. 44
2.4 Anti-Phase Resonator Designs ........................................ 45
  2.4.1 Single-Axis, Two Degrees of Freedom ..................... 45
  2.4.2 Double-Axis, Eight Degrees of Freedom .................. 47
2.5 Multi-Mass Coupling Structures for Tailored Mode Ordering ..... 49
  2.5.1 Internal Spring Coupling .................................... 49
  2.5.2 External Lever Coupling .................................... 50
2.6 Internal Lever Coupling ............................................ 54
  2.6.1 Concept .................................................... 55
  2.6.2 Analytical Model ........................................... 56
  2.6.3 Experimental Results ....................................... 63
2.7 Fabrication-Induced Asymmetry .................................... 64
  2.7.1 Mode-Mixing ................................................ 65
  2.7.2 Asymmetry Quantification ................................... 67
2.8 Common-mode Rejection ............................................. 72
  2.8.1 Analytical Expression ...................................... 73
  2.8.2 G-sensitivity through Applied Vibration ................... 74
  2.8.3 Experimental Results ....................................... 75
2.9 Conclusion ......................................................... 77

3 Quadruple Mass Gyroscope ............................................ 79
  3.1 Introduction ..................................................... 79
  3.2 Structural Design ................................................ 80
    3.2.1 Shuttles for Axis Isolation ............................... 80
    3.2.2 Capacitive Electrodes .................................... 81
    3.2.3 Assumptions ............................................... 82
    3.2.4 Design Progression ....................................... 83
    3.2.5 Modal Resonance Frequencies .............................. 93
  3.3 Quality Factor and the Primary Mechanisms of Energy Loss .... 97
    3.3.1 Viscous Damping ........................................... 98
    3.3.2 Substrate Losses .......................................... 99
    3.3.3 Asymmetry Losses ......................................... 101
    3.3.4 Thermoelastic Damping .................................... 103
    3.3.5 Electronic Losses ........................................ 109
  3.4 Experimental Confirmation of Parameters ....................... 116
    3.4.1 Frequency Responses of Modal Resonances ............... 116
    3.4.2 Q-factor of Modal Resonances ............................ 120
    3.4.3 Parasitic Motion of the Device .......................... 123
  3.5 Instrumentation and Performance ................................ 127
    3.5.1 Front-End Electronics .................................... 127
    3.5.2 Control Loops ............................................. 129
    3.5.3 Bias and Scale Factor .................................... 131
    3.5.4 Frequency Stability and Phase Drift Compensation .... 133
    3.5.5 Acceleration Sensitivity and Compensation ............. 136
  3.6 Conclusion ....................................................... 139
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.2.2</td>
<td>Stress Isolation</td>
<td>204</td>
</tr>
<tr>
<td>6.2.3</td>
<td>Vacuum Sealing</td>
<td>204</td>
</tr>
<tr>
<td>6.3</td>
<td>Influences of Die Attachment</td>
<td>206</td>
</tr>
<tr>
<td>6.3.1</td>
<td>Quality factor</td>
<td>206</td>
</tr>
<tr>
<td>6.3.2</td>
<td>Temperature Range and Capacitive Drifts</td>
<td>207</td>
</tr>
<tr>
<td>6.3.3</td>
<td>Thermal Coefficient of Frequency and Modal Mismatch</td>
<td>208</td>
</tr>
<tr>
<td>6.4</td>
<td>Die Attachment</td>
<td>209</td>
</tr>
<tr>
<td>6.4.1</td>
<td>Epoxy versus Solder</td>
<td>210</td>
</tr>
<tr>
<td>6.4.2</td>
<td>Die Attachment Procedure</td>
<td>211</td>
</tr>
<tr>
<td>6.5</td>
<td>Identification of Packaging Stress</td>
<td>213</td>
</tr>
<tr>
<td>6.5.1</td>
<td>Experimental Procedure</td>
<td>213</td>
</tr>
<tr>
<td>6.5.2</td>
<td>Modeled versus Experimental Results</td>
<td>217</td>
</tr>
<tr>
<td>6.5.3</td>
<td>Reduction of Attachment Area</td>
<td>222</td>
</tr>
<tr>
<td>6.6</td>
<td>Vacuum Sealing</td>
<td>227</td>
</tr>
<tr>
<td>6.6.1</td>
<td>Vacuum Sealing Procedure</td>
<td>228</td>
</tr>
<tr>
<td>6.6.2</td>
<td>Influence on Q-factor</td>
<td>229</td>
</tr>
<tr>
<td>6.7</td>
<td>Conclusion</td>
<td>231</td>
</tr>
<tr>
<td>7</td>
<td>Conclusion</td>
<td>233</td>
</tr>
<tr>
<td>7.1</td>
<td>Contribution of this Dissertation</td>
<td>233</td>
</tr>
<tr>
<td>7.2</td>
<td>Future Research Directions</td>
<td>237</td>
</tr>
<tr>
<td>A</td>
<td>Fundamentals of MEMS Inertial Sensors</td>
<td>265</td>
</tr>
<tr>
<td>A.1</td>
<td>Introduction</td>
<td>265</td>
</tr>
<tr>
<td>A.2</td>
<td>Resonators</td>
<td>265</td>
</tr>
<tr>
<td>A.2.1</td>
<td>Material Properties of Silicon</td>
<td>266</td>
</tr>
<tr>
<td>A.2.2</td>
<td>Spring and Lever Design</td>
<td>269</td>
</tr>
<tr>
<td>A.2.3</td>
<td>Sources of Damping</td>
<td>273</td>
</tr>
<tr>
<td>A.2.4</td>
<td>Electrostatic Actuation</td>
<td>278</td>
</tr>
<tr>
<td>A.2.5</td>
<td>Electrostatic Frequency Tuning</td>
<td>281</td>
</tr>
<tr>
<td>A.2.6</td>
<td>Capacitive Detection</td>
<td>282</td>
</tr>
<tr>
<td>A.2.7</td>
<td>Electromechanical Amplitude Modulation (EAM)</td>
<td>285</td>
</tr>
<tr>
<td>A.2.8</td>
<td>Front-End Electronics</td>
<td>288</td>
</tr>
<tr>
<td>A.2.9</td>
<td>Self-Resonance Electronics</td>
<td>296</td>
</tr>
<tr>
<td>A.2.10</td>
<td>Phase and Amplitude Control</td>
<td>299</td>
</tr>
<tr>
<td>A.2.11</td>
<td>Analog-to-Digital Converters</td>
<td>300</td>
</tr>
<tr>
<td>A.3</td>
<td>Influence of Noise on Inertial MEMS</td>
<td>303</td>
</tr>
<tr>
<td>A.3.1</td>
<td>Noise and Drift Characteristics</td>
<td>303</td>
</tr>
<tr>
<td>A.3.2</td>
<td>Precision and Accuracy of AM and FM Outputs</td>
<td>307</td>
</tr>
<tr>
<td>A.3.3</td>
<td>Linear Drift</td>
<td>308</td>
</tr>
<tr>
<td>A.3.4</td>
<td>Thermomechanical White Noise</td>
<td>308</td>
</tr>
<tr>
<td>A.3.5</td>
<td>Quantization White Noise</td>
<td>310</td>
</tr>
<tr>
<td>A.3.6</td>
<td>Noise in Frequency Modulated Outputs</td>
<td>311</td>
</tr>
<tr>
<td>A.4</td>
<td>Gyroscopes</td>
<td>313</td>
</tr>
<tr>
<td>A.4.1</td>
<td>Newton’s Equations of Motion in a Non-Inertial Frame</td>
<td>313</td>
</tr>
</tbody>
</table>
A.4.2  Ideal Transducer Dynamics for Slow Rotations ($\Omega \ll 2\omega_n$) . . . . 315
A.4.3  Stiffness and Damping Asymmetry . . . . . . . . . . . . . . . . . . . 317
A.4.4  Rigid Body Rotation and Force Misalignment . . . . . . . . . . . . 321
A.4.5  Mass Decoupling and Angle Gain . . . . . . . . . . . . . . . . . . . . 325

B  Silicon Device Fabrication  329
   B.1  Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 329
       B.1.1  Fabrication Approaches . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 329
       B.1.2  Bulk Micromachining Methods . . . . . . . . . . . . . . . . . . . . . . . . . . 330
   B.2  Silicon-on-Insulator Fabrication Process . . . . . . . . . . . . . . . . . . . . 331
       B.2.1  Process Overview . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 332
       B.2.2  Backside Metal Deposition . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 334
       B.2.3  Lithography on Metal . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 337
       B.2.4  Backside Metal Etching . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 339
       B.2.5  Front-side Processing (optional) . . . . . . . . . . . . . . . . . . . . . . . . . . 341
       B.2.6  Lithography on Silicon Dioxide . . . . . . . . . . . . . . . . . . . . . . . . . . . 343
       B.2.7  Silicon Dioxide Hard-Mask Etching . . . . . . . . . . . . . . . . . . . . . . . . 346
       B.2.8  Silicon Device Etching . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 349
       B.2.9  Wafer Dicing . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 357
       B.2.10  Hydrofluoric Acid Release . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 361
       B.2.11  Die Attachment . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 365
       B.2.12  Wire Bonding . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 372
       B.2.13  Vacuum Sealing . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 375
       B.2.14  Challenges . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 378
   B.3  Silicon-on-Glass Fabrication Process . . . . . . . . . . . . . . . . . . . . . . . . 380
       B.3.1  Conventional Silicon-on-Glass Fabrication . . . . . . . . . . . . . . . . . . . 380
       B.3.2  Experimental Process Overview . . . . . . . . . . . . . . . . . . . . . . . . . . . 381
       B.3.3  Challenges . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 385
   B.4  Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 386

C  Protocols  387
   C.1  Silicon-on-Insulator Fabrication . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 388
       C.1.1  Sensor Design and Mask Fabrication . . . . . . . . . . . . . . . . . . . . . . . . . 388
       C.1.2  Backside Metal Deposition and Patterning . . . . . . . . . . . . . . . . . . . . 389
       C.1.3  Front-side Metal Deposition and Etching (optional) . . . . . . . . . . . . . . 390
       C.1.4  Front-side Patterning and Etching . . . . . . . . . . . . . . . . . . . . . . . . . . 391
       C.1.5  Wafer Dicing . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 391
       C.1.6  Scanning Electron Microscopy (optional) . . . . . . . . . . . . . . . . . . . . . . 392
       C.1.7  Die Cleaning and Release . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 392
       C.1.8  Infrared Microscope Inspection . . . . . . . . . . . . . . . . . . . . . . . . . . . . 393
   C.2  Packaging of Silicon Dies . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 395
       C.2.1  Turn On Equipment . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 395
       C.2.2  Bake-out the Chamber (optional) . . . . . . . . . . . . . . . . . . . . . . . . . . . . 396
       C.2.3  Back-out Packages . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 397
       C.2.4  Die Attachment . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 398
List of Figures

1.1 GPS-assisted dead reckoning. .................................................. 2
1.2 Applications for inertial sensors: Consumer electronics (left), automotive (center), inertial navigation (right) [20]. .................................................. 3
1.3 Bias drift. ............................................................................. 11
1.4 Scale factor drift. ................................................................. 11
1.5 Scale factor versus bias for current state of the art commercial MEMS accelerometers. .......................................................... 22
1.6 Scale factor versus bias for current state of the art commercial gyroscopes. .................................................. 23
1.7 Size, weight, and power versus bias for current state of the art commercial gyroscopes. .......................................................... 23
1.8 Academic approaches to increasing inertial sensor performance while reducing SWaP. .......................................................... 26

2.1 Three mode shapes of a disk resonator with central support: translational mode (left), wineglass mode (n=2) (center), wineglass mode (n=3) (right). ........................................................................ 37
2.2 Schematic of a lumped parameter resonator, where stiffness and mass are independently designed. ................................................. 38
2.3 Schematic of a single-axis tuning fork. ........................................ 46
2.4 Schematic of a double-axis tuning fork. .......................................... 48
2.5 Schematic of a double-axis tuning fork with external levers. ................ 51
2.6 Frequency spectrum of a modeled double-axis tuning fork with external levers. .................................................. 54
2.7 Frequency spectrum of a modeled double-axis tuning fork with both external levers and internal lever coupling. .................................................. 56
2.8 Schematic of the lever coupling (top), and simplified mass/spring system (bottom). .................................................. 58
2.9 One-half of the coupling structure with variables defined. .................................................. 58
2.10 Clamped-clamped beam of one-half of the coupling structure and mode shape created from forces and moments applied during the anti-phase resonance of the resonator. .................................................. 59
2.11 Clamped-clamped beam of one-half of the coupling structure and mode shape created from forces and moments applied during the in-phase resonance of the resonator. .................................................. 59
2.12 Plot of the value of the equivalent coupling spring ($k_C$) versus $L/g$ for $h = 1000\mu m$ and $w = 100\mu m$. .................................................. 62
2.13 Plot of the value of the equivalent coupling spring ($k_c$) and the in-phase stiffness ($k_{IN}$) of the coupling structure versus $L/g$ for $h = 1000\mu m$, $w = 100\mu m$, and $g = 10\mu m$. .................................................. 62
2.14 Conceptual schematic of a single axis of the double-axis tuning fork dynamic model. Fabrication imperfections are included in all anchoring springs, $k_A$, and masses are coupled with springs, $k_C$, and levers, $k_L$. .................................................. 67

2.15 Experimental sweep, driving M1 and sensing M2, along with a model of this motion for an ideal device. Small imperfections ignored. .................................................. 68

2.16 Comparison between the ideal model (driving M1 and sensing M2) of the device against the same model with an exaggerated 20% increase in stiffness on a single mass ($\Delta k_1 = 1.2 \cdot k_A$). .................................................. 69

2.17 Direction of change in magnitude of each resonance (Rn) versus location of imperfection ($\Delta k_n$), actuating mass (An), and sensing mass (Sn). Minimally, three sweeps are needed to identify all four imperfections (one orange, green, and blue). 70

2.18 Imperfections identified from a single device, showing 1.4% horizontal and 0.4% vertical change in stiffness across the device, with lower stiffness towards the perimeter of the wafer. .................................................. 71

2.19 Frequency spectrum of two resonators actuated into a constant amplitude of vibration, with application of an external 100 Hz excitation along the axis of vibration. Magnitude of sidebands are correlated to g-sensitivity. ............... 76

3.1 Summary of QMG design progression. .................................................. 84

3.2 Complete Phase 0 device. .................................................. 90

3.3 Single mass of Phase 0 device. .................................................. 90

3.4 Coupling and inner electrodes between left masses of Phase 0 device. .................. 90

3.5 Complete Phase 1 device. .................................................. 91

3.6 Single mass of Phase 1 device. .................................................. 91

3.7 Coupling and inner electrodes between left masses of Phase 1 device. .................. 91

3.8 Complete Phase 2 device. .................................................. 92

3.9 Single mass of Phase 2 device. .................................................. 92

3.10 Coupling and inner electrodes between left masses of Phase 2 device. .................. 92

3.11 Finite element modeling of Phase 0 QMG with spring coupling. .................. 94

3.12 Finite element modeling of Phase 1 QMG with lever coupling. .................. 94

3.13 Comparison of the in-phase / anti-phase mode separation between the Phase 0 and Phase 1 QMG. .................................................. 95

3.14 Finite element modeling of additional, higher order mode shapes of the Phase 1 QMG. .................................................. 96

3.15 Model of two-mass gyroscope used for finite element modeling of anchor loss, consisting of the device, substrate, and perfectly matched layer (PML) for the purpose of absorbing strain energy. .................................................. 100

3.16 Substrate damping versus number of mesh elements for the in-phase and anti-phase mode shapes of Figure 3.15. .................................................. 100

3.17 Demonstration of thermoelastic energy loss due to the interaction between the mechanical and thermal energy domains. .................................................. 104

3.18 Q-factor due to thermoelastic damping, $Q_{TED}$, versus mechanical resonance frequency for three common resonator materials. Beam thickness is 10 $\mu$m. .................. 106

3.19 Q-factor due to thermoelastic damping, $Q_{TED}$, versus mechanical resonance frequency for three different silicon beam widths. .................................................. 107

3.20 Stress and thermal distribution of a two-mass gyroscope. .................................................. 108
3.21 Thermoelastic damping versus number of mesh elements for the in-phase and
anti-phase mode shapes of Figure 3.20. .................................................. 108
3.22 Q-factor versus the amplitude of the carrier voltage for a tuning fork resonator
with transimpedance readout electronics with fit using Equation 3.12. Depicts
curves for three electronics that have four different feedback resistors: 23.7 Ω
(blue), 46.4 Ω (green), 90.9 Ω (red), and 196 Ω (teal). .......................... 114
3.24 Experimental frequency sweep of a Phase 1 QMG by forcing and detecting from
a single mass (left), phase of this sweep (bottom right), and active electrode
diagram (top right). ................................................................. 117
3.25 Experimental frequency sweep of a Phase 1 QMG by forcing a single mass and
detecting all four masses for anti-phase motion (left), phase of this sweep (bottom
right), and active electrode diagram (top right). .................................. 118
3.26 Experimental frequency sweep of a Phase 1 QMG by forcing a single mass and
detecting all four masses for double anti-phase motion (left), phase of this sweep
(bottom right), and active electrode diagram (top right). ......................... 118
3.27 Experimental frequency sweep of a Phase 1 QMG by forcing a single mass and
detecting all four masses for opposing in-phase motion (left), phase of this sweep
(bottom right), and active electrode diagram (top right). ......................... 119
3.28 Experimental frequency sweep of a Phase 1 QMG by forcing a single mass and
detecting all four masses for in-phase motion (left), phase of this sweep (bottom
right), and active electrode diagram (top right). .................................. 119
3.29 Amplitude decay measurement for anti-phase Phase 1 QMG motion after given
an initial impulse for three different pressure levels. The rate of amplitude decay
is directly related to the damping in the system (τ or Q-factor) .................. 121
3.30 Q-factor versus pressure for the anti-phase resonance of a Phase 1 QMG (left),
along with magnitude versus frequency for higher pressure levels not shown in
Figure 3.29 (right). ................................................................. 122
3.32 Demonstration of out-of-plane detection using a laser doppler vibrometer.
Photograph of experimental setup (top left), along with out-of-plane displacement
of a Phase 1 QMG versus frequency of an in-plane actuation force applied to a
single mass (right). ................................................................. 125
3.33 In-plane and out-of-plane displacement of a Phase 1 QMG at various locations
on the die. The device was actuated using the in-plane electrode on the x-axis
of the device. ................................................................. 126
3.34 Photograph of the PCB front-end electronics used for signal conditioning of the
actuation and detection signals prior to interfacing the MEMS, which is also
inserted into the PCB. ................................................................. 129
3.35 Schematic of the open-loop gyroscope control scheme. .......................... 130
3.36 Example of an Allan variance curve of the zero rate output of the gyroscope. 132
3.37 Allan variance curve of the zero rate output of the gyroscope. − 1/2 and + 1/2 fits
are shown as dashed red lines. .......................................................... 134
3.38 Allan variance of the gyroscope frequency, as determined by the phase-locked
control loop. − 1/2 fit is shown as a dashed red line. ............................ 135
3.39 Rate red noise versus frequency white noise. ........................................ 136
3.40 Gyroscope bias versus static acceleration, before and after compensation using the quadrature and frequency of the device (left). Linear compensation was applied from both frequency (top left) and quadrature (bottom left).

4.1 Schematic of an amplitude modulated device (left), along with its voltage output and response (right). The response is the envelope of the voltage output.

4.2 Schematic of a frequency modulated device (left), along with its voltage output and response (right). The response is the frequency of the voltage output.

4.3 Complete FM accelerometer with high-frequency and linear scale factor.

4.4 Single mass of a high-frequency FM accelerometer with linear scale factor.

4.5 Complete FM accelerometer with low-frequency and linear scale factor.

4.6 Single mass of a low-frequency FM accelerometer with linear scale factor.

4.7 Electrode placement on single mass of an FM accelerometer. The electrode types are mirrored on the second mass.

4.8 Design comparison between an FM accelerometer with linear (left) and non-linear (right) scale factor, as dictated by mode ordering.

4.9 Finite element modeling of both the low-frequency (left) and high-frequency (right) FM accelerometer designs.

4.10 Description of how the in-phase and anti-phase resonances of the FM accelerometer interact for acceleration detection.

4.11 Schematic of a dual tuning fork FM accelerometer and control architecture of both temperature, T, and acceleration, A, measurement.

4.12 Example of the differential frequency output of a dual tuning fork FM accelerometer, and the robustness of this output due to temperature.

4.13 Output of the dual FM accelerometer over time during an intentional temperature ramp from 70 °C to 30 °C. The device has a Q-factor of 600 with individual tines plotted in blue and red and the differential output plotted in black.

4.14 Allan deviation of the dual FM accelerometer over integration time with room temperature fluctuations. The device has a Q-factor of 350,000 with plots for both an individual tine and the differential output.

4.15 Frequency shift of a linear FM accelerometer versus acceleration with different tuning voltages, revealing a linear tunable scale factor for the device with maximum magnitude of ~ 2,000 ppm/g.

4.16 Frequency shift of a non-linear FM accelerometer versus acceleration, revealing a maximum scale factor of ~ 20,000 ppm/g.

4.17 Amplitude versus time of a low-frequency FM accelerometer after being given an impulse. Decay reveals a τ of 1300 s and Q-factor of 2.3 million.

5.1 Schematic of the control architecture to be used for scale factor calibration using the co-fabricated gyroscope and torsional resonator. Final output and scale factor error are shown.

5.2 Process flow for the described co-fabricated bulk micro-machined gyroscope and calibration platform.

5.3 Schematic of the signal routing used for the vibratory gyroscope, along with a wire bonding diagram for pulling the signals off-chip.

5.4 Model of the final co-fabricated device. The four primary layers include the metal electrodes for signal routing, the gyroscope device layer, the torsional calibration platform, and the substrate.
5.5 Complete device. ................................................................. 182
5.6 Routing metal. ................................................................. 182
5.7 Gyroscope layer. ............................................................... 182
5.8 Torsional platform. ............................................................ 182
5.9 Bonding metal. ................................................................. 182
5.10 Substrate. ....................................................................... 182
5.11 Photographs of wafers partially through the fabrication sequence for the gyro-
scope and torsional platform, before and after wafer bonding. ............... 183
5.12 Optical profilometry of the silicon torsional resonator after etching. Due to the
large gaps, depth could be monitored between etching to achieve the proper etch
depth of 300 µm. ........................................................................ 184
5.13 Photographs of co-fabricated device, top (left) and bottom (right). [133] . . . 185
5.14 Amplitude versus time of an SOI platform under vacuum, showing a Q-factor of
70,000. ..................................................................................... 187
5.15 Photograph of a packaged SOI platform (left), along with a schematic of the
control of the structure during stability testing (right). ............................... 187
5.16 Measured velocity (left) and frequency (right) stability of the SOI platform de-
picted in Figure 5.15. .................................................................. 188
5.17 Finite element modeling of the torsional platform. ................................. 189
5.18 Depiction of the phase variation in the torsional resonator when analyzed for
out-of-plane motion using a laser doppler vibrometer during in-plane actuation. 190
5.19 In-plane (left) and out-of-plane (right) motion of the torsional resonator during
in-plane actuation. In-plane displacement was measured using optical image
processing and out-of-plane displacement was measured using a laser doppler
vibrometer. .................................................................................. 191
5.20 Macro-scale torsional calibration stage with attached commercial gyroscope. The
attached rate table supplies a unknown input rotation, Ω_{IN}, while the torsional
stage supplies a known calibration signal, Ω_{Cal}. ................................ 193
5.21 Output of the commercial gyroscope when Ω_{IN} = 0.2 Hz and Ω_{Cal} = 10 Hz. . 194
5.22 Output of the commercial gyroscope on the macro-scale torsional stage with
induced temperature ramp (left), along with the output of the device before and
after calibration by the torsional stage (right). ...................................... 195
5.23 Process flow for the proposed co-fabricated bulk micro-machined gyroscope and
calibration platform. ................................................................. 198

6.1 X-ray images of two devices with opposing packaging processes. Both underwent
the same temperature profile, but the left was performed in air, while the right
was performed in vacuum. ............................................................... 207
6.2 Temperature profile of a device during die attachment when following the out-
lined procedure (right), along with photograph of the experimental setup (left). 212
6.3 Testing procedure of measuring device resonance frequency before and after die
attachment. ............................................................................. 214
6.4 Example frequency sweeps of a device before and after packaging. A slightly
higher frequency can be observed after packaging. ............................... 215
6.5 Packaging procedure and X-ray image of die attachment area. The length and
width of the area was measured and compared to the shift in resonance frequency. 216
6.6 X-ray images of devices after die attachment, showing the random reflow of solder during attachment. .......................................................... 216
6.7 Finite element model of the packaging process. The silicon device, along with the ceramic package, was modeled with variable die attachment area. The anti-phase resonance of the device was then measured before and after a large static temperature shift. Due to symmetry of the device and resonance, only one quarter of the device was modeled. ........................................ 218
6.8 Modeled and experimental results showing the resonance frequency shift versus die attachment diameter. .................................................. 219
6.9 Critical die attachment diameters observed in the model of frequency shift before and after die attachment. Each corresponds to when the die attachment reaches a location where the resonant device is anchored to the substrate. ............. 220
6.10 Experimental results showing the thermal coefficient of frequency of several devices with varying die attachment diameter. .................. 221
6.11 Influence of die attachment area on deflection of springs after die attachment. Full die attachment imparts stress causing the hairpin springs to deflect (left). Reduced die attachment size alleviates stress and eliminates deflection (right) . 223
6.12 Influence of liquid hydrofluoric acid on adhesion strength of backside metal. The smallest metal pads to survive exposure were approximately 2 mm in diameter. 224
6.13 Vapor-phase hydrofluoric acid release of SOI devices: undercut depth versus exposure time of two different devices (right). Measurements were performed using infrared microscope imaging (left). ................................. 226
6.14 Die after 45 minutes of vapor HF exposure with front-side protected. Perimeter was exposed to the vapor, but the silicon dioxide in the center remained. ........... 227
6.15 Temperature profile of a device during vacuum sealing when following the outlined procedure (right), along with influence of temperature ramp rate on the final Q-factor of the device after sealing (left). ..................................... 228
6.16 Frequency sweeps of a QMG packaged with the temperature profile shown in Figure 6.15. Under vacuum, the Q-factor improves from 20 to 1000 in both modes of resonance. ......................................................... 230

7.1 Prototype single-chip system, comprised of two accelerometers and one high performance z-axis gyroscope. ................................. 240
7.2 Proposed approach for creating a 6-axis IMU (right) through the use of multiple 3-axis single-chip systems (left). Redundant accelerometers provide continual compensation for chip alignment. ................................. 241
7.3 Quadruple mass gyroscope with differential detection electrodes applied to each mass. Signals can be combined to isolate each mode shape. ................. 242
7.4 Amendment to the process flow of Figure B.1 to reduce die attachment asymmetry. 245

A.1 The crystalline planes of single crystal silicon, denoted by Miller indices. ....... 267
A.2 Comparison of in-plane elastic properties of standard <100> (blue) and <111> (red) silicon wafers when rotating around the normal vector: Young’s modulus (left), Poisson’s ratio (center), and shear modulus (right). ................................. 269
A.3 A clamped-guided spring of length L, width w, and thickness t (not-shown): undeformed (left) and deformed (right). ................................. 270
A.4 Common springs designs: (a) crab leg, (b) hairpin, (c) double hairpin, (d) mirrored hairpin, and (e) shuttled hairpin. ................................. 271
A.5 A lever design for scaling and possibly inverting displacement. 272
A.6 A lever design for scaling and transforming displacement orthogonally. 272
A.7 Schematic of a damped harmonic oscillator. 274
A.8 Amplitude decay of a mechanical resonator, used to measure damping. 275
A.9 Normalized magnitude and phase versus normalized frequency for various Quality factors. 276
A.10 Schematics of two types of electrostatic actuation: variable area (left) and variable gap (right). 279
A.11 Photograph of lithography-defined comb (left) and parallel plate electrodes (right). 279
A.12 Cross-sectional view of an SOI resonator with single-sided forcing and detection (top), and with an electrical overlay (bottom). 283
A.13 Cross-sectional view of an SOI resonator with differential forcing and detection (top), and with an electrical overlay (bottom). 284
A.14 Spectral density plots demonstrating the electromechanical amplitude modulation technique. 286
A.15 Current detection using both passive (left) and active (right) techniques. 288
A.16 Simplified schematic of an operational amplifier. 289
A.17 Active current detection using a transimpedance amplifier with feedback capacitor. 290
A.18 Poles $s_1$ and $s_2$ in the complex plane as feedback capacitor, $C_f$, increments from 0.1 to 10 pF. The system is critically damped when $s_1 = s_2$. 292
A.19 Magnitude and phase of the system from Figure A.18, with and without the optimal feedback capacitance. 293
A.20 Simplified instrumentation amplifiers displaying the design progression of the classic design, shown in Figure A.21. Single operational amplifier with low input impedance (left) and a design with two additional buffering amplifiers, but fixed gain (right). 294
A.21 Basic schematic of the classic instrumentation amplifier using three separate operational amplifiers, three sets of symmetric resistors, and a programmable gain resistor. 294
A.22 Front-end electronics for combining a differential current output into a common signal. Inset are example current outputs from the MEMS resonator (left) along with output after combination (right). 295
A.23 Front-end electronics for creating a differential AC driving voltage with common DC offset. Inset are example actuation voltages for the input (left) and output (right) versus time. 296
A.24 Input and output sinusoidal signals from a MEMS resonator, displaying the necessary signal processing for self-resonance. 297
A.25 Control scheme for self-resonance implementation using noncoherent AGC. 298
A.26 Control scheme for self-resonance implementation using coherent AGC. 299
A.27 Example of a two bit Flash ADC: Analog input $V_{IN}$ is compared to each voltage on a voltage ladder, the height of which is encoded into a binary signal. 301
A.28 Simplified schematic of a first-order Sigma-Delta ADC. 302
A.29 Simplified schematic of a Voltage-to-Frequency ADC. 302
A.30 Examples of how various classifications of noise and linear drift behave in the time (left) and frequency domains (right). 303
A.31 Allan deviation of a signal with multiple noise sources. In order from left to right: White noise ($\tau^{-1/2}$), pink noise ($\tau^0$), red noise ($\tau^{1/2}$), and linear drift ($\tau^1$). 

A.32 Phase noise of an oscillator versus frequency offset from the carrier. From left to right: Random walk of frequency ($f^{-4}$), flicker of frequency ($f^{-3}$), white frequency noise and random walk of phase ($f^{-2}$), flicker of phase ($f^{-1}$), and white phase noise ($f^0$).

A.33 Point in a rotating, non-inertial reference frame.

A.34 Multi-axis resonator in a rotating, non-inertial reference frame.

A.35 Multi-axis resonator in a local frame of reference with both stiffness and damping asymmetry.

A.36 Simplified multi-axis resonator in a local frame of reference with misaligned principal axes of stiffness and damping.

A.37 Multi-axis resonator with stiffness and force not acting through the center of mass.

A.38 Multi-axis resonator in a local frame of reference with force asymmetry.

A.39 Multi-axis resonator in a local frame of reference with decoupled mass matrix.

The principal axes of mass are defined orthogonal, but allowed to be misaligned to the local frame.

B.1 Process flow for the described silicon-on-insulator (SOI) fabrication process.

B.2 Backside of an SOI wafer after metallization of chrome and gold.

B.3 Backside of an SOI wafer after with patterned photoresist masking prior metallization. The pattern consists of a grid of assorted sizes.

B.4 Backside of an SOI wafer after etching exposed chrome and gold and removal of the photoresist mask.

B.5 Patterned backside metallization using a different lithographic mask, comprising of a grid of 2mm diameter circles.

B.6 Frontside of an SOI wafer after metallization and patterning.

B.7 Sensor lithography after alignment to existing frontside metallization.

B.8 Frontside of an SOI wafer after etching of the silicon dioxide hard-mask: the full wafer (left), along with a focused view on a single device (right).

B.9 Frontside of an SOI wafer after etching of the silicon dioxide hard-mask: portion of a device (left), along with a closer view of the smallest features (right).

B.10 DRIE process flow, used for deep etching of silicon with vertical sidewalls.

B.11 DRIE when a buried oxide layer is reached. Charging of oxide during RIE etches the sidewall passivation, creating excessive silicon at the interface.

B.12 Frontside of an SOI wafer after etching of the silicon device layer: the full wafer (left), along with a focused view on a single device (right).

B.13 Frontside of an SOI wafer after etching of the silicon device layer: portion of a device (left), along with a closer view of the smallest features (right).

B.14 Frontside of an SOI wafer after etching of the silicon device layer: view of the proof mass with etch holes (left), along with a closer view (right).

B.15 Images of a Phase 2 QMG: IR (left) and SEM cross-section (right).

B.16 IR images after incremental etching: +1.75 min (left) and +2.0 min (right).

B.17 IR images after incremental etching: +2.25 min (left) and +2.5 min (right).

B.18 SEM of a cross-section of the device, located at the proof mass, for two different etching times.
B.19 SEM of a cross-section of the device, located at the parallel plate electrodes, for two different etching times.

B.20 SEM of a cross-section of the device, located at the comb electrodes, for two different etching times.

B.21 Frontside of an SOI wafer after device etching and sensor protection with photoresist, prior to dicing.

B.22 Clean silicon dies with surface silicon dioxide, ready for hydrofluoric acid release.

B.23 Stiction of parallel plates. Expanded view of electrode (left) and closer view showing contact between proof mass and electrode (right).

B.24 Idonus vapor-phase HF etching equipment for dry etching of buried silicon dioxide.

B.25 Idonus IR microscope for inspection of buried silicon dioxide etch rate.

B.26 Images from the IR microscope, illustrating the etch rate of the vapor-phase etcher by inspecting a device once per hour of etching.

B.27 Etch rate of buried silicon dioxide for devices from two different fabrication processes: in-house SOI fabrication (blue) and commercial SOI-MUMPS (red).

B.28 Vacuum furnace used for die attachment and getterless vacuum sealing: expanded view of entire setup.

B.29 Two DIP packages on top of a graphite beam, before (left) and after (right) removal of the connection between pins.

B.30 Packages and silicon dies placed in vacuum furnace before package bake-out.

B.31 Die attachment assembly of package. Graphite frame inserted with solder preform (left), insertion of silicon die (center), insertion of graphite lid (right).

B.32 Weights placed on top of the graphite lid to apply pressure to bond.

B.33 Vacuum furnace used for die attachment and getterless vacuum sealing: inside of vacuum champer with one package setup for die attachment and the second with a bimetallic temperature gauge.

B.34 Vacuum furnace used for die attachment and getterless vacuum sealing: view of device and gauge through furnace window.

B.35 Vacuum furnace used for die attachment and getterless vacuum sealing: applied weights to reduce vibrations from vacuum pump.

B.36 Silicon die attached to package after die attachment.

B.37 WestBond Gold/Aluminum Wire Bonder, model 747677E.

B.38 High vacuum furnace from SST International, model 3150.

B.39 Inspex X90 X-ray System: Expanded view.

B.40 Inspex X90 X-ray System: View through window.

B.41 Image of device springs after die attachment.

B.42 Boron dopant concentration in silicon verses resistivity, plotted according to Equation B.1.

B.43 Process flow for the experimental silicon-on-glass (SOG) fabrication process.

B.44 Photograph of the wafer stack after dissolving of the substrate using EDP. Silicon device layer bonded to a glass substrate. [133]

B.45 Photograph of individual dies on the wafer shown in Figure B.44. Photographs taken from the top (left) and through the glass from the bottom (right). [133]

B.46 Microscope photographs of the dies shown in Figure B.45. Photographs taken from the top, showing a debonded anchor sitting upside-down on top of the proof mass (left), as well as from the bottom, showing anchors that remained attached (right). In both images, significant undercutting of anchors can be seen.
C.1 Overview of the vacuum furnace and support equipment. 396
C.2 Two DIP packages on top of a graphite beam, before (left) and after (right) removal of the connection between pins. 398
C.3 Packages and silicon dies placed in vacuum furnace before package bake-out. 399
C.4 Die attachment assembly of package. Graphite frame inserted with solder pre-form (left), insertion of silicon die (center), insertion of graphite lid (right). 400
C.5 Weights placed on top of the graphite lid to apply pressure to bond. 401
C.6 Die attachment assembly inside the vacuum furnace, with bimetallic temperature gauge for measuring package temperature. 401
C.7 Temperature gauge and weights visible through observation window in vacuum furnace. 402
C.8 Additional weights applied to the top of the vacuum furnace for reducing vibrations. 402
C.9 Silicon die attached to package after die attachment. 403
C.10 West-bond wire bonding station. 405
C.11 Breadboard implementation of front-end electronics for capacitive actuation and detection. 410
C.12 Actuation circuit (IA = AD620, R_G = 500kΩ, R_1 = 1kΩ, C_1 = 1µF). 411
C.13 Detection circuit (OA = OP177, IA = AD620, R_G = 500kΩ, R_F = 10kΩ, and C_F = 10pF). 411
C.14 Interface locations for BNCs (V_A, V_DC, V_O) and ports for jumper wires to the socket (V_A1, V_A2, V_D1, V_D2). 414
C.15 Basic set-up of equipment and connections. 414
C.16 Initial settings under the “Modulation” tab. 420
C.17 Initial settings under the “Connectivity” tab, along with a saturated carrier frequency in the “Oscilloscope”. 422
C.18 Initial settings under the “Connectivity” tab, along with an unsaturated carrier frequency. 422
C.19 Resonance sweep of the drive mode. 424
C.20 Spectroscope data during resonance sweep of the drive mode. 425
C.21 Spectroscope data as the phase “Set Point” is adjusted for PLL2 until maximum value is reached. 426
C.22 Spectroscope data as the gyroscope is given a sinusoidal input rotation from a rate table. 428

E.1 “Accel1” - FM Accelerometer Mask 1. 438
E.2 “DieAttach1” - Die Attachment Definition 1. 438
E.3 “DieAttach2” - Die Attachment Definition 2. 439
E.4 “iX” - Commercial Fabrication Run at iX-Factory 1. 439
E.5 “Mask13” - Gyroscope Mask 13. 440
E.6 “Mask13” - Gyroscope Mask 13 - Labels. 441
E.7 “Mask14” - Gyroscope Mask 14. 441
E.8 “Mask15” - Gyroscope Mask 15. 442
E.9 “Mask16” - Gyroscope Mask 16. 443
E.10 “Mask17” - Gyroscope Mask 17. 443
# List of Tables

1.1 Physical phenomena utilized for rotation measurement. ........................................ 5
1.2 Resonator materials used in Coriolis vibratory gyroscopes. ................................ 7
1.3 Design approaches for Coriolis gyroscopes. ......................................................... 9
1.4 Gyroscope performance grades. Adapted from [40]. ............................................... 17
1.5 Accelerometer performance grades. [40] ............................................................... 18
1.6 Performance of commercial sensors in terms of scale factor stability, bias repeata-
    bility, and SWaP. ......................................................................................... 24

2.1 Double-axis tuning fork mode shapes and properties. .............................................. 48

3.1 Summary of QMG design progression. ................................................................. 85
3.2 Electronic damping fitting parameters for Figure 3.23. ........................................... 115

4.1 Amplitude modulated versus frequency modulated detection schemes. .................. 144
4.2 Improvement in dynamic range of FM accelerometer as a function of Q-factor. ....... 165

6.1 Material properties of silicon and Kyocera A440 [148]. ........................................ 217

A.1 Physical properties of the individual crystalline planes of silicon [151], [152], [153]. 269
A.2 Physical properties of silicon. ............................................................................... 269
A.3 Power laws for various types of noise and drift with respect to power spectral
density, Allan variance, and phase noise (only applicable to frequency measurements). The
top group is applicable to any signal, while frequency measurements also include the bottom segment. ................................................................. 304

B.1 SOI wafer properties. ......................................................................................... 333
B.2 Properties of photoresist AZ4620. ...................................................................... 344
B.3 Silicon dioxide etching recipe for the STS Advanced Oxide Etcher. .................... 347
B.4 Photoresist stripping recipe for the Matrix System One Stripper. ......................... 349
B.5 Deep silicon etching recipe for the Plasma-Therm FDRIE. .................................. 350
B.6 Aluminum wire properties for wirebonding. ....................................................... 373
B.7 Wirebonder settings for bonding to gold and silicon bond pads of standard and
deep packages. ......................................................................................... 373
ACKNOWLEDGEMENTS

“If I have seen further it is by standing on the shoulders of giants”

-Isaac Newton, 1676

First, I would like to thank Professor Andrei Shkel for his inspiration and technical expertise in the field of inertial MEMS, both of which have shined a light on the many challenges and unanswered questions still present; I am personally indebted to him for this insight and have appreciated his time and patience. Dr. Alexander Trusov has also been instrumental in this process, for which he has my eternal gratitude. I would also like to thank Professor Faryar Jabbari for his time and commitment to being a part of this committee and reviewing this work.

I would also like to show my appreciation for the entire MicroSystems Laboratory at the University of California, Irvine; I do not believe there has been a member for which I do not owe some useful insight, along with fostering a friendly atmosphere. More specifically, I would like to thank Dr. Sergei Zotov for his guidance; Doruk Senkal, Sina Askari, and Dr. Igor Prikhodko for their numerous technical discussions; as well as Dr. Christopher Raum and Montgomery Rivers for their knowledge of MEMS fabrication.

This work would have also not been possible without UCI’s Integrated Nanosystems Research Facility (INRF), UCLA’s Nanoelectronics Research Facility (Nanolab), as well as the many vendors of Appendix D. Specifically, I would like to thank Jake Hess for his time and effort, as well as Mohammad Kebaili for his fabrication assistance.

Finally, I would also like to thank the generous funding agencies that have supported this work throughout the years, specifically the Defense Advanced Research Projects Agency (DARPA), the Office of Naval Research (ONR), and the Naval Surface Warfare Center Dahlgren Division (NSWCDD). This work has been partially supported by DARPA grant N66001-12-C-4035, as well as ONR and NSWCDD grants N00014-09-1-0424 and N00014-11-1-0483.
CURRICULUM VITAE

Brenton Ross Simon

EDUCATION

Ph.D. in Mechanical and Aerospace Engineering 2014
University of California, Irvine  Irvine, CA

M.S. in Mechanical and Aerospace Engineering 2009
University of California, Davis  Davis, CA

B.S. in Mechanical Engineering 2007
University of Cincinnati  Cincinnati, OH

EXPERIENCE

Graduate Student Researcher 2009-2014
University of California, Irvine  Irvine, CA

Graduate Student Researcher 2007-2009
University of California, Davis  Davis, CA

Undergraduate Student Researcher 2007
University of Cincinnati  Cincinnati, OH

Cooperative Education Student 2007
Pedco E&A Services  Cincinnati, OH

Cooperative Education Student 2006
Siemens VDO  Wuerzburg, Germany

Cooperative Education Student 2004
General Electric  Cincinnati, OH

Cooperative Education Student 2003
Cinergy Corporation  Cincinnati, OH
PUBLICATIONS


PATENTS


Abstract of the Dissertation

Mode Ordering in Anti-Phase Driven MEMS Gyroscopes and Accelerometers

By

Brenton Ross Simon

Doctor of Philosophy in Mechanical and Aerospace Engineering

University of California, Irvine, 2014

Professor Andrei M. Shkel, Chair

Inertial sensors have a wide spectrum of applications, ranging from consumer electronics to precision navigation. As these devices continue to maximize performance, while minimizing Size, Weight, and Power (SWaP) requirements, these opportunities will only continue to expand. Vibratory Micro Electro Mechanical System (MEMS) inertial sensors are uniquely positioned in this landscape due to their low SWaP metrics and high potential for enhanced performance. In this dissertation, the fundamental challenges behind the further advancement of these devices are explored, with a number of potential solutions proposed.

For vibratory gyroscopes, one of these fundamental challenges is the tradeoff between rate and acceleration sensitivity; both of which are enhanced by low frequencies of operation. Anti-phase resonances are typically employed to decouple this influence; however, when conventional flexures are used, the anti-phase vibratory mode is forced to a higher frequency, reducing rate sensitivity. For this reason, a novel coupling structure has been designed, analyzed, modeled, fabricated, and tested. This structure is experimentally
shown to selectively stiffen in-phase vibration, creating a high degree of modal separation in excess of 120%, a value that is believed to be the highest in published literature, along with reducing acceleration sensitivity by over 20 fold. Theoretical analysis shows that the observed frequency separation can continue to be expanded with this technique, which is only limited by fabrication constraints. This type of structure was also applied to a new frequency modulated accelerometer, and shown to enhance the mechanical scale factor by over 20 times.

Resonator quality factor is another critical element that can be maximized to enhance the performance of some inertial sensors. By identifying the primary energy loss mechanisms within the frequency range of interest, each mechanism was modeled and minimized through design and fabrication. The result of this work was a resonator with quality factor of 2.34 million and decay constant of 1300 s, both of which are also believed to be the highest in published literature for microfabricated structures. In addition to these highlights, investigations also include an in-run scale factor calibration method, through the use of an integrated torsional rate stage, as well as packaging considerations for enhanced temperature robustness.
Chapter 1

Introduction

1.1 Motivation

Reliable navigation has been a consistent pursuit of humanity since nearly the dawn of recorded history. The desire for safe and efficient travel and trade has lead to the development of a wide range of technologies over the millennia: landmarks, celestial observation, and sea currents have been some of the more popular options throughout civilization, with the global positioning system (GPS) supplanting all of these options in modern times [19]. Each of these options, however, have one common theme: they rely upon an external point of reference. Due to this requirement, each of these technologies can become obscured due to poor conditions, or even malicious jamming from a third party. In order to avoid this vulnerability, a common pursuit is to precisely measure the motion of the navigating object and use this information to track the path of the object from a known starting position. This concept is known as inertial navigation.

Long-term inertial navigation can be considered the ultimate goal of inertial sensor development, with sensor performance becoming increasingly stringent as operation time increases. An example of this can be seen in the table of Figure 1.1, which shows the
- GPS-guided dead reckoning:
  - GPS signal determines location
  - When GPS signal unobtainable, inertial sensors estimate location
- Error builds up over time

<table>
<thead>
<tr>
<th>Resolution</th>
<th>1 sec</th>
<th>1 min</th>
<th>1 hr</th>
<th>1 day</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 μg</td>
<td>5 μm</td>
<td>18 cm</td>
<td>64 m</td>
<td>37 km</td>
</tr>
<tr>
<td>100 μg</td>
<td>0.5 mm</td>
<td>1.8 m</td>
<td>6.4 km</td>
<td>3700 km</td>
</tr>
</tbody>
</table>

Figure 1.1: GPS-assisted dead reckoning.

Error accumulation of two accelerometers with different resolutions over time. As will be later shown, accelerometers with even these degrees of resolution are difficult to achieve, which typically restricts accurate inertial navigation to only a few minutes of operation.

To overcome these challenges, modern navigation does not solely rely on any one given technology, but rather multiple in the form of an inertial navigating system (INS). By optimally combining the data from multiple sources through the use of a Kalman filter, the location of the object can be tracked precisely and with redundancy. In addition to inertial sensors, these inputs also typically include GPS, radar, and magnetic sensors.

An example of how these various technologies can operate in synergy is shown in Figure 1.1, which combines inertial navigation with GPS; while the continuous inertial data accumulates error over time, the intermittent GPS signal is able to reposition the object to bound the error growth.
While long-term navigation may be the most performance-demanding application for inertial sensors, additional opportunities continue to arise as these devices increase in performance and reduce its cost, size, weight, and power (C-SWaP). A few examples include low C-SWaP sensors for consumer electronics, and robust devices for automotive safety and stability, Figure 1.2. Each of these applications can involve drastically different sensor requirements with unique design and control elements.

1.2 Background

There are many different design philosophies that can be used for the creation of inertial sensors. In the mid 20th century, many of the highest performance devices were based on precision machining of complex structures; however, that has since changed as technology has advanced. Listed here are the most common physical phenomena used for detecting rotation, along with materials and designs of Coriolis-type devices, which are of particular interest to this work.

1.2.1 Principles of Operation

Under the most fundamental of classifications, the phenomena used by gyroscopes to detect rotation can be divided into two categories: 1) Cross-axis energy coupling, and 2)
Differential frequency measurement. Cross-axis force measurements are typically limited in resolution by how accurately the amplitude of the sensor can be detected, while differential frequency measurements are limited by time resolution. There have been many physical phenomena that have been shown to be capable producing one of these two types of measurements [21]-[24]; however, the three most popular include the Coriolis effect, Sagnac interference, and Larmor precession [19]. The Coriolis effect is force-based, while Sagnac interference and Larmor precession rely on frequency measurements.

Coriolis and Sagnac-type devices currently dominating the market in the form of vibratory and optical gyroscopes; however, the nuclear magnetic resonance gyroscope (NMRG) is a promising new technology, but has yet to produce any available products to date. Coriolis devices function by inducing a standing vibrational wave within a two-axis oscillator. When the device experiences a rotation, some of the energy of the forced vibrational wave is transmitted to the degenerate axis. This effect can then be used to measure rate, by directly quantifying this energy transfer, or angle, by allowing the standing wave to freely rotate. Because this type of detection uses a physical vibration, the energy consumed by the device is a function of the mass and damping of the vibrating structure, both of which lead to low power consumption. Rate detection, however, is based off of the amplitude of a vibrating structure, which leads to poor environmental robustness.

Optical devices utilizing Sagnac interference function by taking a beam of light with a precise frequency and splitting it, sending both components around the parameter of an enclosed area in opposite directions. Because both beams initially have an identical
frequency and light travels at a fixed velocity, any drift in the phase between the two beams after traveling around the enclosed area can be correlated to a rotation. This technique requires a macro-scale fabrication approach, leading to higher costs; however, because there are no moving parts, it maintains a high environmental robustness.

While there is still considerable research to be performed for all types of inertial sensors, the current market for Coriolis-type devices tend to be for low C-SWaP applications, where high performance is desirable, but not critical. Sagnac-type devices, such as the fiber optic gyroscope (FOG) and ring laser gyroscope (RLG), currently dominate for high-performance applications, where a higher C-SWaP is still tolerable. A short summary of these trends is shown in Table 1.1. These differences will be discussed further, later in this chapter.

### 1.2.2 MEMS Gyroscope Materials

The main focus of this work is on Coriolis-type devices. When designing such a device there are many considerations that must be taken into account, one of which is the material of the vibrating structure. Ever since the first monolithic microfabricated gyroscope in 1991 [25], which was later refined at a tuning-fork structure in 1994 [26], silicon has
been a popular choice for the resonator material. In many respects, this has been due to the wide variety of fabrication techniques available for their construction, compatibility with integrated circuit (IC) fabrication, as well as a high Young’s modulus. Despite these advantages, silicon has a high stiffness sensitivity to temperature, as can be seen in Table A.2, which directly influences the frequency of such resonators and leads to thermal stabilization requirements.

While most commercial silicon gyroscopes have some type of temperature calibration to reduce this thermal influence, a second option was found in replacing the resonator material. Single-crystal quartz is a uniquely capable resonator material with a number of interesting properties. First, particular cuts of quartz crystal experience less than $200 \text{ ppm}$ variation in their stiffness over the common environmental temperature range of $-55$ to $125^\circ C$ [27], [28]. Second, the material is piezoelectric, allowing transduction of the material directly through electric fields. Finally, like silicon, the crystalline structure of the material allows for selective wet etching of certain crystal faces, potentially simplifying fabrication. However, the main benefit is the temperature stability of the material, making it still the most popular material for high stability resonator fabrication to date [29]. Despite the success of this material for timing applications, the lack of intricate fabrication options has lead to an inability to fabricate devices of increased complexity, such as gyroscopes and accelerometers.

In order to meet this need, while still maintaining some of the desirable qualities of the material, an amorphous form of quartz was later used by companies such as Northrop Grumman and Sagem, known as fused quartz or fused silica [30]. While the
<table>
<thead>
<tr>
<th>Material</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silicon</td>
<td>Well known fabrication, Dry or wet etching, Doping possible</td>
<td>Large temperature drifts</td>
<td>Draper</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tuning Fork [25], [26]</td>
<td></td>
</tr>
<tr>
<td>Quartz</td>
<td>Wet etching possible, Small temperature drifts, Piezoelectric</td>
<td>Limited dry etching capabilities</td>
<td>Quartz</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tuning Fork [27]-[29]</td>
<td></td>
</tr>
<tr>
<td>Fused Silica</td>
<td>Material properties independent of orientation, low internal damping</td>
<td>Difficult / expensive to fabricate</td>
<td>HRG [30]</td>
</tr>
</tbody>
</table>

Table 1.2: Resonator materials used in Coriolis vibratory gyroscopes.

temperature insensitivity of the material was lost due to the amorphous state, it still maintained low internal damping characteristics and now had properties independent of a preexisting crystal orientation. Northrop Grumman used this material to create the hemispherical resonator gyroscope (HRG), which currently holds the record for Coriolis vibratory gyroscope (CVG) bias stability and white noise. The primary consequence of using this material, though, is limited dry etching capabilities, which is common to current commercial microfabrication processes.

A summary of the advantages and challenges of each of these common materials is provided in Table 1.2.

1.2.3 Types of Vibratory Gyroscopes

Due to fabrication limitations, the chosen resonator material has a strong impact on the design space of the sensor. Putting these limitations aside for the moment, there are five general design strategies for Coriolis vibratory gyroscopes (CVGs): 1) Bulk Acoustic Wave, 2) Tuning Fork, 3) Disk or Ring, and 4) Wineglass. A brief summary of these options is provided in Table 1.3.
The first gyroscopes used for inertial navigation were macro-scale spinning rotor gyroscopes. These devices were highly complex with precision machined components, both of which lead to a relatively high cost and rate of mechanical failure. This eventually allowed optical gyroscopes to replace these devices with their increased performance and higher reliability. While there is still some research into using electrostatic levitation to create a spinning rotor gyroscope on the micro-scale [31], this is not an aggressively advancing field, with most research focusing on monolithic vibratory gyroscopes instead.

One of the main challenges of vibratory gyroscopes is sensitivity to environmental disturbances, such as shock, vibration, and thermal variations. Bulk acoustic wave (BAW) gyroscopes attempt to solve this issue by no longer relying on a low frequency suspended mass, but rather a high frequency bulk-mode resonance. Quartz is typically used for these devices, but any piezoelectric material is acceptable. Other than this difference, they function in a similar way as other types of Coriolis vibratory gyroscopes [32]; however, due to their high frequency and low amplitude of motion, they are typically low performance devices. The main advantage of piezoelectric gyroscopes reside in their environmental robustness, potentially allowing them to be an optimal choice for certain harsh conditions.

Low frequency vibratory gyroscopes also attempt to mitigate the effects of external vibration through the use of tuning fork structures [26]. Due to the complex designs of such devices, they are generally fabricated from silicon and null common-mode impulses through the use of an anti-phase resonance. While such a design approach is ideal, the presence of fabrication imperfections deteriorates the effectiveness of this technique, as
<table>
<thead>
<tr>
<th>Type</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bulk Acoustic</td>
<td>Low cost, High shock survivability</td>
<td>Low performance</td>
<td>BAW Gyroscopes [23], [24]</td>
</tr>
<tr>
<td>Tuning Fork</td>
<td>Mod.-High performance, Customizable design</td>
<td>High external vibration sensitivity</td>
<td>Draper Tuning Fork [26]</td>
</tr>
<tr>
<td>Disk / Ring</td>
<td>Mod.-High performance, Low vibrational sensitivity, Innate structural symmetry</td>
<td>Low amplitude of motion</td>
<td>Disk / Ring Gyroscopes [33]-[35]</td>
</tr>
<tr>
<td>Wineglass</td>
<td>Very high performance</td>
<td>High cost</td>
<td>HRG [30]</td>
</tr>
</tbody>
</table>

Table 1.3: Design approaches for Coriolis gyroscopes.

will be discussed in Sections 2.7 and 3.3.3. The primary benefit of this design, though, is that a low frequency and high amplitude of motion can be maintained, thus maximizing performance.

Disk or ring gyroscopes somewhat occupy the design space between tuning fork and bulk gyroscopes; their performance and robustness lays somewhere between the two design strategies [33], [34], [35]. Due to their symmetric design, these devices have an additional benefit of typically low frequency mismatch between the required degenerate vibratory modes, which can also aid performance, despite their relatively higher frequency and low amplitude when compared to tuning fork devices; however, this is at the cost of a low angle gain. Once again, due to the complex geometry of such devices, they are typically fabricated from silicon.

The final type of Coriolis vibratory gyroscope are 3-D wineglass structures. These devices operate similarly to disk gyroscopes, but instead of relying on an internally etched, 2-D geometry to lower the resonance frequency to acceptable levels, wineglasses are able
to reduce the stiffness by expanding into the third dimension. By maintaining a solid
dome structure, it is believed there are improvements to both frequency symmetry and
Q-factor. Currently, there are only a few commercially available gyroscopes of this form,
e.g. Northrop Grumman and Sagem [30], each requiring high cost macro-scale fabrication
techniques.

1.3 Trends and Challenges

Coriolis vibratory gyroscopes are capable of extremely high performance, as demonstrated
by the HRG, yet also small sizes and low power requirements, as shown by most consumer
sensors. In addition, increasing the robustness of these devices to non-inertial and off-axis inertial effects is desired. Combining these properties is the primary focus of ongoing research.

In this section, standard terminology will first be defined, followed by discussions on
device stability and environmental robustness to establish the primary challenges of this work.

1.3.1 Terminology

While some performance metrics of inertial sensors appear universally known and ac-
cepted, there still appears to be some nomenclature discrepancy between various aca-
demic groups and commercial manufacturers. For this reason, terminology will be defined
below, separated into four sections: 1) Measurement, 2) Uncertainty, 3) Input conditions,
and 4) Applications and environmental conditions. While a far majority of these terms
are being borrowed from the IEEE for consistency [36], there are few that have been
included which the IEEE has not defined, yet are useful when discussing certain applications. The testing of many of these parameters are well defined by the IEEE [37], [38].

Measurement

- **Zero Offset**: The gyro output when the input rate is zero, generally expressed as an equivalent input rate. It excludes outputs due to hysteresis and acceleration.

- **Bias \( (\circ/\text{hr}) \)**: The average over a specified time of gyro output measured at specified operating conditions that has no correlation with input rotation or acceleration.

- **Scale Factor**: The ratio of a change in output to a change in the input intended to be measured. Scale factor is generally evaluated as the slope of the straight line that can be fitted by the method of least squares to input-output data.

Uncertainty
• Random Walk: A zero-mean Gaussian stochastic process with stationary independent increments and with standard deviation that grows as the square root of time.

• Angle Random Walk (°/√h). The angular error buildup with time that is due to white noise in angular rate. This error is typically expressed in degrees per square root of hour or degrees per hour per square root hertz. The conversion is:

\[ \frac{\circ}{\sqrt{Hz}} = \frac{\circ}{\sqrt{h}} \cdot \frac{1}{60}. \]

• Rate Random Walk ((°/h)/√h). The drift rate error buildup with time that is due to white noise in angular acceleration.

• Stability: A measure of the ability of a specific mechanism or performance coefficient to remain invariant when continuously exposed to a fixed operating condition.

• Bias Instability (°/h): The random variation in bias as computed over specified finite sample time and averaging time intervals. This nonstationary (evolutionary) process is characterized by a 1/f power spectral density. It is typically expressed in degrees per hour.

• Bias Repeatability (°/h): As bias, but also between multiple runs under specified environmental conditions. This term is not defined by the IEEE, but is used in industry.

• Bias Asymmetry (°/h): The difference between the bias for positive and negative inputs, typically expressed in degrees per hour.

• Scale Factor Asymmetry: The difference between the scale factor measured with positive input and that measured with negative input, specified as a fraction of the
scale factor measured over the input range. Scale factor asymmetry implies that
the slope of the input-output function is discontinuous at zero input. It must be
distinguished from other nonlinearities.

• Hysteresis Error: The maximum separation due to hysteresis between upscale-
going and down-scale-going indications of the measured variable (during a full-range
traverse, unless otherwise specified) after transients have decayed. It is generally
expressed as an equivalent input.

• Linearity Error: The deviation of the output from a least-squares linear fit of the
input-output data. It is generally expressed as a percentage of full scale, or percent
of output, or both.

• Drift Rate: The component of gyro output that is functionally independent of input
rotation. It is expressed as an angular rate.

• Systematic Drift Rate: The component of drift rate comprising bias, environment-
tally sensitive drift rate, and elastic-restraint drift rate.

• Random Drift Rate: The random time-varying component of drift rate. Multiple
contributing factors include thermal cycling / stress changes, mass loading through
contaminates, and material changes due to diffusion or radiation [39].

• Quantization Noise: The random variation in the digitized output signal due to
sampling and quantizing a continuous signal with a finite word length conversion.
The resulting incremental error sequence is a uniformly distributed random variable
over the interval 1/2 least significant bit (LSB).
• Resolution: The largest value of the minimum change in input, for inputs greater than the noise level, that produces a change in output equal to some specified percentage (at least 50%) of the change in output expected using the nominal scale factor.

Input Conditions

• Dead Band: A region between the input limits within which variations in the input produce output changes of less than 10% (or other small value) of those expected based on the nominal scale factor.

• Dynamic Range ($dB$): The ratio of the input range to the resolution.

• Full Range: The algebraic difference between the upper and lower values of the input range.

• Bandwidth ($Hz$): Change in input rotation which degrades performance by 3 dB. Usually measured by applying a sinusoidal input rotation and increasing the frequency until the required output error is reached. This term is not defined by the IEEE, but is used in industry.

Applications and Environmental Conditions

• Cross-coupling Errors: The errors in the gyro output resulting from gyro sensitivity to inputs about axes normal to an input reference axis.

• Sensitivity: The ratio of a change in output to a change in an undesirable or secondary input. For example: a scale factor temperature sensitivity of a gyro or accelerometer is the ratio of change in scale factor to a change in temperature.
- **Turn-on Time (s):** The time from the initial application of power until a sensor produces a specified useful output, though not necessarily at the accuracy of full specification performance.

- **Warm-up Time (s):** The time from the initial application of power for a sensor to reach specified performance under specified operating conditions.

- **Run-down Time (s):** The time interval after removal of excitation during which either (1) the gyro maintains quoted performance metrics or (2) the sensing element motion decays to a specified amplitude.

- **C-SWaP ($\$, m$^3$, kg, W):** Shorthand for cost, size, weight, and power. These are the primary logistical factors which influence the marketability of a device and areas in which MEMS inertial sensors have a competitive advantage over competing technologies. This term is not defined by the IEEE, but is used in industry.

- **Temperature Range ($\Delta^\circ$C):** A range of temperatures for which the device will operated with the listed performance. This term is not defined by the IEEE, but is used in industry.

- **Maximum Temperature ($^\circ$C):** The maximum temperature the device can experience before damage. Usually this is measured when the device is not currently in operation. This term is not defined by the IEEE, but is used in industry.

- **Maximum Shock (g):** The maximum acceleration the device can experience before damage. Usually this is measured when the device is not currently in operation. This term is not defined by the IEEE, but is used in industry.
1.3.2 Stability

The stability of bias and scale factor over time are two of the most critical parameters when determining sensor performance. While instability can arise in a number of forms, as mentioned in the previous section, the culmination of these effects determine the resolution of the device and effectiveness at identifying an inertial input. There are a number of natural phenomena that influence sensor stability, ranging from drifts in the mechanical properties of the resonators, to variable phase delays in the control electronics; stability typically refers to the innate properties of the sensor itself and not environmental conditions. Identifying and minimizing these phenomena is crucial for enhancing sensor performance.

Due to the importance of device stability, inertial sensors are generally classified for certain applications based upon these ideal performance characteristics. While some variability exists between these categories, which is dependant upon the source, four common gyroscope grades include: 1) Rate, 2) Industrial, 3) Tactical, and 4) Inertial.

One of the greatest challenges of these devices is that most gyroscopes output the rotational rate of the sensor, yet most applications desire an angle output. To meet this need, the output of a gyroscope is typically integrated to form the angle of rotation. The challenge with this approach is that small errors in the rate output of a gyroscope accumulate over the course of the integration, leading to large drifts in angle over time. For this reason, the gyroscope rate errors must be tightly controlled when an output angle is desired.

When comparing the performance of the three gyroscope grades, rate grade gyro-
Table 1.4: Gyroscope performance grades. Adapted from [40].

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Rate</th>
<th>Industrial</th>
<th>Tactical</th>
<th>Inertial</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angle Random Walk (ARW)</td>
<td>&gt; 0.5</td>
<td>&gt; 0.5</td>
<td>0.5-0.05</td>
<td>&lt; 0.001</td>
<td>$^\circ/\sqrt{hr}$</td>
</tr>
<tr>
<td>Bias Stability</td>
<td>10-1000</td>
<td>10-100</td>
<td>0.1-10</td>
<td>&lt; 0.01</td>
<td>$^\circ/hr$</td>
</tr>
<tr>
<td>Scale Factor Accuracy</td>
<td>100-1000</td>
<td>100-500</td>
<td>10-100</td>
<td>&lt; 1 ppm</td>
<td>ppm</td>
</tr>
<tr>
<td>Full Scale Range</td>
<td>50-1000</td>
<td>&gt; 500</td>
<td>&gt; 500</td>
<td>&gt; 400</td>
<td>$^\circ/sec$</td>
</tr>
<tr>
<td>Maximum Shock in 1 ms</td>
<td>$10^3$</td>
<td>$10^3$</td>
<td>$10^3-10^4$</td>
<td>$10^3$</td>
<td>g</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>&gt; 70</td>
<td>∼ 100</td>
<td>∼ 100</td>
<td>∼ 100</td>
<td>Hz</td>
</tr>
</tbody>
</table>

scopes have the worst of the three. These types of sensors are typically used when only a rate output is desired, alleviating the need to integrate the result to obtain the angle of rotation, thus allowing the user additional measurement tolerance. Inertial grade gyroscopes, on the other hand, require the greatest performance due to this error accumulation. There is no lower bound to device performance for this application, because further stability enhancements simply extend the time of accurate angle output before recalibration is required. Tactical grade sensor performance lays between these two extremes. Like inertial grade sensors, the output of tactical grade devices is typically integrated to form an angle measurement; however, unlike inertial grade applications, the angle output is typically bounded. A few examples of applications of these devices include platform stability control (such as for cameras or antennas), down-hole drilling, and GPS-aided navigation. A detailed summary of the performance metrics for devices within each of these regimes is provided in Table 1.4, along with a visual representation of each category in Figure 1.2.

A similar grading system exists for accelerometers; however, is a little less refined. Automotive applications, for both safety and stability, have dominated the field of MEMS accelerometers since their conception and have guided design goals. With their continued
increases in performance, navigational applications have also arisen. The requirements for each of these applications is provided in Table 1.5.

### 1.3.3 Environmental Robustness

In addition to the native stability of Coriolis vibratory gyroscopes, environmental robustness is a second major concern. Environmental sensitivities can come in many forms: temperature, acceleration, vibration, and shock are a few of the more critical influences. Sensors generally have ratings on the maximum range of these parameters to which the sensor can be exposed and ensure the quoted performance. This is typically achieved through on-chip calibration or by quoting parameters for which the user must apply compensation to the recorded data. In addition to these sensitivities, there are also maximum rating for which the device can not exceed or risk damage. This may be due to temperature in the form of melting or cracking the packaging, or acceleration causing contact of the vibratory structure with the substrate or electrodes, damaging the resonator.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Automotive</th>
<th>Navigation</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range</td>
<td>±50 (airbag)</td>
<td>±1</td>
<td>g</td>
</tr>
<tr>
<td></td>
<td>±2 (stability system)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency Range</td>
<td>DC-400 Hz</td>
<td>DC-100 Hz</td>
<td></td>
</tr>
<tr>
<td>Resolution</td>
<td>&lt; 100 (airbag)</td>
<td>&lt; 0.004</td>
<td>mg</td>
</tr>
<tr>
<td></td>
<td>&lt; 10 (stability system)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Off-axis Sensitivity</td>
<td>&lt; 5</td>
<td>&lt; 0.1</td>
<td>%</td>
</tr>
<tr>
<td>Nonlinearity</td>
<td>&lt; 2</td>
<td>&lt; 0.1</td>
<td>%</td>
</tr>
<tr>
<td>Max. Shock in 1 ms</td>
<td>&gt; 2000</td>
<td>&gt; 10</td>
<td>g</td>
</tr>
<tr>
<td>Temperature Range</td>
<td>-40 to 85</td>
<td>-40 to 80</td>
<td>°C</td>
</tr>
<tr>
<td>TC of Offset</td>
<td>&lt; 60</td>
<td>&lt; 0.05</td>
<td>mg/°C</td>
</tr>
<tr>
<td>TC of Sensitivity</td>
<td>&lt; 900</td>
<td>±50</td>
<td>ppm/°C</td>
</tr>
</tbody>
</table>

Table 1.5: Accelerometer performance grades. [40]
As mentioned previously, the stiffness of silicon has a high sensitivity to temperature, which directly influences the resonance frequency of such devices. While this effect can greatly reduce performance if left uncompensated, it typically does not cause any permanent damage. When attempting to compensate such a sensor using an on-chip thermometer, there are some challenges with respect to the resolution of the thermal measurement, as well as thermal variability between the resonator and thermometer; however, such techniques can go a long way for improving performance. In some respects, a larger challenge arises in acceleration sensitivity.

Low frequencies of motion are beneficial to the performance of both gyroscopes and accelerometers by directly enhancing the mechanical scale factor of these devices. For gyroscopes, this creates issues when exposed to vibrations near the operational frequency, by causing non-Coriolis induced vibrations in the sense mode of the sensor, thus creating false measurements. This issue becomes an increasing concern for high-Q devices by further exacerbating the effects of the input. Not only can impulses at the resonance frequency of the device effect the output, but the sensor is also sensitive to static and lower frequency acceleration. For tuning fork devices, the in-phase mode of resonance determines how well the structure resists deformation of these quasi-static forces. Even with differential detection, capacitive nonlinearities can influence the output of the sensor due to changes in the gap width, which is especially apparent when parallel plate electrodes are used. Even when these displacements are small with respect to the driven amplitude of motion, they can still be quite large when compared to the displacement induced by the Coriolis force. Accelerometers have a similar tradeoff when choosing a
resonance frequency, but instead of low operational frequencies reducing environmental robustness, the bandwidth of the device is reduced.

These environmental challenges make it advantageous to design devices for specific applications, especially when low C-SWaP is desired.

1.4 MEMS Inertial Sensor Prior Art

MEMS inertial sensors have been present both academically and commercially for over 20 years. While a significant amount of progress has been made in this time, there are still many questions that remain to be answered. In this section, a review of the current state of commercially available devices is presented, along with the current academic approaches to further push the limits of performance.

1.4.1 Commercial Devices

MEMS inertial sensors began to appear in commercial products in the mid-1990s, with their first widespread use in automotive airbag sensors [41]. Since then, their popularity has only continued to increase as applications have expanded, both due to demand, as well as technological advancement. This advancement has been in the form of both increases in performance, resulting in higher accuracy measurements, as well as reductions in Cost, Size, Weight, and Power (C-SWaP), which reduces the influence of the sensor on the measured object. Both of these factors are key for expanding the breadth of the sensor landscape.

A compilation of some of the larger companies known to commercially produce MEMS gyroscopes include: Analog Devices, Silicon Sensing, Systron Donner (BEI), Bosch Semi-
conductors, Memsense, Honeywell, InvenSense, STMicrotechnologies, and Sagem [42].

A similar list for accelerometers include: Analog Devices, Colibrys, Kionix, Endevco, InvenSense, Bosch Semiconductors, Memsense, Honeywell, STMicrotechnologies, and Northrop Grumman [42].

In order to compare these commercial devices for an accurate representation of the available commercial landscape, three parameters were chosen for comparison: Bias repeatability, scale factor stability, and SWaP. Bias repeatability was chosen as an estimation of the bias error, because it assumes that the sensor will not be calibrated before every run. Many companies prefer to quote bias instability, because this is the minimum achievable bias for the device given long-term averaging and ideal conditions; however, this is generally not applicable for most applications. Scale factor stability is also compared as the second uncertainty component to device performance. Finally, a single quantity for SWaP is determined using the volume, weight, and power requirements of each device. Because these SWaP metrics are not typically reported for single-axis devices, only data gathered from 6-axis inertial measurement units (IMUs) were used for these comparisons (the only exception is the HRG, which is quoted as a single-axis sensor). A list of the metrics from each device is provided in Table 1.6.

Figure 1.5 provides a plot of scale factor stability versus bias repeatability for the MEMS accelerometers examined. While there is definitely some variability in performance, multiple devices have currently achieved bias repeatabilities of 1 mg, along with scale factor stabilities of 300 ppm. A similar plot is provided in Figure 1.6 for gyroscopes, which does not only include devices based on vibratory MEMS devices, but also optical
Figure 1.5: Scale factor versus bias for current state of the art commercial MEMS accelerometers.

This plots shows the spectrum of performance of each type of device, and while it appears to be not flattering to MEMS gyroscopes, it is important to consider the SWaP metrics for most applications. Figure 1.7 displaces the cumulative SWaP of each device versus bias repeatability. In this figure, it is much easier to observe the advantages of MEMS sensors, which beats competing technologies in SWaP metrics by several orders of magnitude, despite a minor cost in performance.
Figure 1.6: Scale factor versus bias for current state of the art commercial gyroscopes.

Figure 1.7: Size, weight, and power versus bias for current state of the art commercial gyroscopes.
<table>
<thead>
<tr>
<th>Label</th>
<th>Company</th>
<th>Product</th>
<th>Type</th>
<th>Accelerometer</th>
<th>Gyroscope</th>
<th>Packaging</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF (ppm)</td>
<td>SF Bias (°/hr)</td>
<td>SF Bias (mg)</td>
</tr>
<tr>
<td>A</td>
<td>Goodrich</td>
<td>MinIM</td>
<td>MEMS (Si)</td>
<td>1800</td>
<td>20</td>
<td>100</td>
</tr>
<tr>
<td>B</td>
<td>Goodrich</td>
<td>SiIMU02</td>
<td>MEMS (Si)</td>
<td>1500</td>
<td>10</td>
<td>500</td>
</tr>
<tr>
<td>C</td>
<td>Goodrich</td>
<td>SiIMU04</td>
<td>MEMS (Si)</td>
<td>150000</td>
<td>100</td>
<td>3000</td>
</tr>
<tr>
<td>D</td>
<td>Honeywell</td>
<td>OEM-HG1900</td>
<td>MEMS (Si)</td>
<td>200</td>
<td>1</td>
<td>150</td>
</tr>
<tr>
<td>E</td>
<td>Honeywell</td>
<td>OEM-HG1930</td>
<td>MEMS (Si)</td>
<td>300</td>
<td>5</td>
<td>300</td>
</tr>
<tr>
<td>F</td>
<td>Honeywell</td>
<td>UIMU-HG1700-AG58</td>
<td>Optic (RLG)</td>
<td>500</td>
<td>1</td>
<td>150</td>
</tr>
<tr>
<td>G</td>
<td>Kearfott</td>
<td>KI-4901S</td>
<td>Optic (RLG)</td>
<td>100</td>
<td>0.1</td>
<td>50</td>
</tr>
<tr>
<td>H</td>
<td>KVH</td>
<td>CNS-5000</td>
<td>Optic (FOG)</td>
<td>9000</td>
<td>7.5</td>
<td>1000</td>
</tr>
<tr>
<td>I</td>
<td>KVH</td>
<td>CG-5100</td>
<td>Optic (FOG)</td>
<td>4000</td>
<td>50</td>
<td>1000</td>
</tr>
<tr>
<td>J</td>
<td>KVH</td>
<td>TG-6000</td>
<td>Optic (FOG)</td>
<td>6000</td>
<td>10</td>
<td>3000</td>
</tr>
<tr>
<td>K</td>
<td>Northrop Grumman</td>
<td>LITEF uIMU-M</td>
<td>MEMS (Si)</td>
<td>1500</td>
<td>3</td>
<td>1400</td>
</tr>
<tr>
<td>L</td>
<td>Northrop Grumman</td>
<td>LN-200 FOG</td>
<td>Optic (FOG)</td>
<td>300</td>
<td>0.3</td>
<td>100</td>
</tr>
<tr>
<td>M</td>
<td>Sensonor</td>
<td>STIM300</td>
<td>MEMS (Si)</td>
<td>300</td>
<td>0.75</td>
<td>500</td>
</tr>
<tr>
<td>N</td>
<td>Systron Donner</td>
<td>MMQA-200-100</td>
<td>MEMS (Qtz)</td>
<td>5000</td>
<td>3</td>
<td>5000</td>
</tr>
<tr>
<td>O</td>
<td>Systron Donner</td>
<td>SDI500-AB00 Quartz</td>
<td>MEMS (Qtz)</td>
<td>200</td>
<td>1</td>
<td>200</td>
</tr>
<tr>
<td>P</td>
<td>Systron Donner</td>
<td>SDI500-BB00 Quartz</td>
<td>MEMS (Qtz)</td>
<td>300</td>
<td>1.5</td>
<td>300</td>
</tr>
<tr>
<td>Q</td>
<td>Systron Donner</td>
<td>SDI500-CB00 Quartz</td>
<td>MEMS (Qtz)</td>
<td>300</td>
<td>2</td>
<td>400</td>
</tr>
<tr>
<td>R</td>
<td>Colibrys</td>
<td>MS9001</td>
<td>MEMS (Si)</td>
<td>300</td>
<td>0.75</td>
<td>N/A</td>
</tr>
<tr>
<td>S</td>
<td>Colibrys</td>
<td>MS9002</td>
<td>MEMS (Si)</td>
<td>300</td>
<td>1.5</td>
<td>N/A</td>
</tr>
<tr>
<td>T</td>
<td>Honeywell</td>
<td>Q-Flex QA-750</td>
<td>MEMS (Qtz)</td>
<td>1000</td>
<td>8</td>
<td>N/A</td>
</tr>
<tr>
<td>U</td>
<td>Honeywell</td>
<td>GG1320AN</td>
<td>Optic (RLG)</td>
<td>N/A</td>
<td>N/A</td>
<td>5</td>
</tr>
<tr>
<td>V</td>
<td>KVH</td>
<td>DSP-1750 Analog</td>
<td>Optic (FOG)</td>
<td>N/A</td>
<td>N/A</td>
<td>500</td>
</tr>
<tr>
<td>W</td>
<td>KVH</td>
<td>DSP-1760</td>
<td>Optic (FOG)</td>
<td>N/A</td>
<td>N/A</td>
<td>200</td>
</tr>
<tr>
<td>X</td>
<td>KVH</td>
<td>DSP-3400</td>
<td>Optic (FOG)</td>
<td>N/A</td>
<td>N/A</td>
<td>1500</td>
</tr>
<tr>
<td>Y</td>
<td>Northrop Grumman</td>
<td>HRG Scalable SIRU</td>
<td>HRG</td>
<td>N/A</td>
<td>N/A</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 1.6: Performance of commercial sensors in terms of scale factor stability, bias repeatability, and SWaP.
1.4.2 Academic Publications

There are two primary paths towards the advancement of inertial sensors: 1) Improving performance, and 2) Reducing SWaP. As can be seen in Figure 1.7, existing technologies are located at the top and right on a plot of SWaP versus performance. Reducing both of these factors open up new applications for these devices, therefore shifting the various approaches represented on this graph to the lower left is the major goal of academic research.

Figure 1.8 illustrates the necessary shift in performance metrics and can be used to represent the four major approaches of inertial sensor research: 1) Miniaturize the hemispherical resonator gyroscope, 2) Continue increasing the performance of optical gyroscopes, as well as further miniaturization, 3) Increase the performance of MEMS inertial sensors, and 4) Develop new technology that has been enabled by recent advancements in synergetic fields. The current state of each approach is discussed below.

**Approach 1) Miniaturization of the hemispherical resonator gyroscope (HRG).** The hemispherical resonator gyroscope produced by the Northrop Grumman corporation has been shown to produce remarkable performance, largely accredited to its high degree of symmetry [30]. This symmetry is a result of both intrinsic properties of its modal resonances, as well as post-fabrication trimming of mass and stiffness. As a macro-scale device, a larger number of options exist for these fabrication and trimming processes; however, at the cost of C-SWaP of the final device. Due to these high costs, yet excellent performance, these devices have only found use in space applications. Should similar performance be obtained after the conversion to smaller devices, capable
of batch-fabrication, a great number of additional applications would present themselves.

A number of approaches are being taken to complete this miniaturization, particularly at academic institutions, investigating new fabrication techniques for the 3-D resonant structure. A few of the more unique attempts include conformal deposition onto an isotropically etched silicon mold [65], [66], [67], or sphere [68], and wafer-level glass blowing using furnaces [69], or blow torches [70]. Each of these methods have been shown to produce 3-D hemispherical resonators, with work continuing towards achieving gyroscope performance metrics.

**Approach 2) Continue increasing the performance of optical gyroscopes, as well as further miniaturization.** Current commercial optical gyroscopes (RLGs
and FOGs) are macro-scale devices that have been continually increasing in performance and reducing in SWaP. While this is possible by continually refining manufacturing processes, it results in gradual advancement that will eventually begin tapering off due to fundamental challenges, such as the trade off between device diameter and performance [71]. For this reason, many researches are choosing to pursue batch-fabricated solutions that still utilize the Sagnac effect for the detection of rotation. The most promising approach to this goal are resonant micro optical gyroscopes (RMOGs) [72].

RMOGs are similar to RLGs and FOGs; however, they are able to fundamentally decouple the laser source from the sensing element, reducing the influence of any noise imparted from the laser. It accomplishes this by using the laser to actuate two counter-propagating beams in a high-Q ring resonator. The difference between the wavelengths of these two beams can then be detected and related to rotation. These devices have been made possible by the rapid development of photonic integrated circuits (PICs) [73], and are estimated to be eventually comparable in performance to their high SWaP RLG and FOG counterparts, but with C-SWaP metrics comparable to traditional MEMS gyroscopes [72]. While these devices are still in their academic infancy, prototypes have so far achieved bias stabilities of 400 °/hr [74].

**Approach 3) Increase the performance of MEMS inertial sensors.** While MEMS inertial sensors excel in terms of C-SWaP, allowing them to be used in a variety of applications, their performance has lagged in comparison to competing technology. For this reason, there has been considerable research toward enhancing the performance of these devices by examining and enhancing the individual components. These com-
ponents include raising the mechanical Q-factor of the transducer [75], making design modifications to the resonant structure for reducing bias [76], including specific electrodes for canceling quadrature [77], implementing new control algorithms, such as mode reversal [78] and active mode matching [79], or expanding the functionality of devices to multi-axis detection [80].

This approach will serve as the focus for this work, enhancing performance through the rejection of common-mode acceleration, maximization of Q-factor, implementation of novel detection schemes, exploring on-chip calibration, and reducing thermal effects due to packaging.

**Approach 4a) Develop new technology: Nuclear Magnetic Resonance Gyroscopes (NMRG).** Nuclear magnetic resonance (NMR) was first discovered by Isidor Rabi in 1938 [81], for which he was later awarded the 1944 Nobel Prize in physic [82]. NMR is a physical phenomenon in which the quantum state of the nucleus of certain atom can be influenced by electromagnetic radiation, when in the presence of a magnetic field. Because the technique is non-destructive by only temporarily manipulating the quantum spin of certain atoms, it is widely used for identifying organic substances (NMR spectroscopy), as well as advanced medical imaging (magnetic resonance imaging (MRI)). Another, less common use for this technology is inertial sensing [83].

NMR gyroscopes utilize a small cloud of known vapor for detecting rotation. The composition of this vapor can vary, but generally consists of at least two different isotopes: an alkali vapor and a noble gas. Due to their single valance shell electron, alkali atoms are easily excited using NMR, while the noble gas atoms are not. This is beneficial because
the alkali atoms can serve as an actuation force for the nobel gas atoms, transferring their quantum spin to the nobel gas atoms through random collisions in the vapor cloud. The nobel gas atoms are then not only robust to external magnetic fields, but also can be chosen with higher atomic weights to enhance sensitivity.

The process for using such a vapor cloud for detecting rotation first involves applying a static magnetic field across the cloud and optically pumping the alkali atoms to a fixed energy state, both parallel to the $z$-axis of the device. This polarizes the alkali atoms, which is transferred to the nobel gas atoms through collisions. An AC magnetic field is then applied perpendicularly to the static field, creating a net magnetic field that precesses around the $z$-axis. The frequency of this precession is chosen to be the Larmor frequency of the nobel gas, which is substantially lower than the Larmor frequency of the alkali atoms.

Due to the polarization of the nobel gas, a third magnetic field is generated that is precessing slowly when compared to that of the alkali atoms. This causes the alkali atoms to precess at their Larmor frequency around the combination of all three magnetic fields. The alkali atom can then be monitored and the precession of the nobel gas extracted through modulation. Because the nobel gas precesses at a fixed frequency which is determined by the properties of the gas and applied magnetic field, any change in frequency that is observed can be attributed to external rotation around the $z$-axis [83], [84].

Many of the theoretical issues concerning the functionality of NMRGs had been solved in the 1960s and 70s. As a result of this research, several macro-scale devices were
produced, capable of producing bias stabilities of less than 0.1 °/hr [83]. Since then, there
have been considerable advancements in chip-scale atomic clocks that can be directly
applied to the miniaturization of NMRGs. Northrop Grumman has shown developments
of such miniaturized devices with experimental noise levels of less than 0.12 °/√hr [85].

**Approach 4b) Develop new technology: Atom Interferometry (Cold Atom Sensors).** Another approach to utilizing atoms for inertial measurement is present through the use of atom interferometry [86], sometimes known as cold atom sensors. This approach is similar to that of current optical devices (RLGs and FOGs) in that it utilizes a Sagnac-like effect for detecting rotation; however, instead of using photons to conduct the measurement, it utilizes atoms. Due to the wave-particle duality of matter, atoms move with a wavelength that is inversely proportional to their momentum, which is also known as their de Broglie wavelength. Much like the photons of optical sensors, the phase difference in the de Broglie wavelength of atoms can also be observed and correlated to rate.

As described by Bouyer [86], these devices operate by optically trapping a small cloud of atoms in a vacuum, and reducing the temperature of this cloud to sub-µK temperatures using laser cooling. The atom cloud is then ejected, typically upward, and passes through a series of three Raman beams, each separated by a fixed interval of time, which alter the quantum state of the cloud. Because the atom cloud has been supercooled, they are assumed to be initially in their ground state. The first beam is designed to excite half of the atoms to their next discrete state of energy. While the unexcited half of the cloud continues on their initial trajectory, the excited half splits to form a new trajectory due
to the momentum transfer from the beam. There are now two discrete atom trajectories that are departing away from one another, when the clouds experience the second beam. This beam mirrors their current trajectory, and instead of the clouds diverging, they are now heading towards one another. When they collide, a third beam recombines the atom cloud and the phase different can be measured. Through the separation and recombination of the atom clouds, both groups would travel the same distance given static conditions; however, not in the presence of rotation, due to a Sagnac-like effect [87].

Recently, atom interferometry has been used to measure both rotation and acceleration with bias stability of $0.04 \, ^\circ/hr$ and $6.7 \times 10^{-12} \, g$, respectively; however, this is accomplished by cooling a cloud of Rubidium isotopes to $3 \, nK$ and ejecting them into a $10 \, m$ high vacuum chamber [88]. While smaller, more portable devices are in development for inertial navigation applications [89] which are able to isolate phase changes from rotation and acceleration [90], a majority of efforts appear to be focused on enhancing sensitivity by extending run time [91]. Large atom interferometers are believed to be able to push the limits of inertial detection, serving as important cosmological tools for detecting gravitational waves [92].

1.5 Research Objective

The primary objective of this research is to enhance the performance of MEMS inertial sensors by improving the fundamental limits of stability, as well as enhancing environmental robustness. Because MEMS already excel in terms of C-SWaP characteristics, improvements in the fundamental and environmental stability of these structures has the
potential to allow them to dominate the inertial sensing market. In order to achieve this, silicon MEMS gyroscopes and accelerometers with a tuning fork structure were designed, packaged, and analyzed.

A number of strategies were used to achieve this goal. First, design modifications were assessed to decrease the sensitivity of the gyroscope to external acceleration (Chapter 2), as well as maximize the resonant Q-factor to enhance the mechanical scale factor (Chapter 3). Similar design modifications were assessed for a MEMS accelerometer, as well, along with the implementation of a frequency modulated detection scheme (Chapter 4). An in-run method of scale factor calibration was also investigated (Chapter 5), along with the influences of device packaging on sensor performance (Chapter 6).

1.6 Dissertation Outline

This dissertation is separated into seven chapters and four supporting appendices. Chapter 1 has supplied a thorough background on the motivation for inertial sensors, along with various design approaches and challenges. The figures of merit have been defined, along with the current state of commercial products, as well as academic research. The objective of this research has also been presented.

Chapter 2 discusses the benefits of utilizing multi-mass resonators as the vibratory elements of Coriolis vibratory gyroscopes, along with the importance of the mode-shape. Anti-phase motion is shown to reject common-mode vibration, as well as prevent energy loss, thus improving quality factor, and hence device performance. Mode ordering of the parasitic mode-shapes is also shown to be a critical element of robust performance, and a new, negative stiffness coupling mechanism is presented which allows modal ordering
and isolation.

Chapter 3 utilizes the design philosophy of Chapter 2 to present a modified quadruple mass gyroscope. Several design iterations are shown and discussed, along with the corresponding improvements in performance. A detailed assessment of the primary energy loss mechanisms of tuning-fork devices is presented, along with analytical representations of this mechanisms and finite element modeling. Finally, the influence of parasitic motion is discussed, along with a method of compensation by using the quadrature signal of the device.

Chapter 4 discusses an alternative approach to traditional amplitude modulated acceleration detection, through acceleration-sensitive frequency modulation. The main benefits of this approach include the reduction of alignment errors between gyroscopes accelerometers through single-chip fabrication, as well as a high frequency resolution by Q-factor maximization. Two design approaches are presented, along with an assessment of trade-offs. Finally, a device is presented with what is believed to be the current record holder of both Q-factor and decay constant for a microfabricated resonator.

Chapter 5 discusses a new fabrication methodology for on-chip scale factor calibration of gyroscopes, through the use of an embedded torsional resonator. The results of an intensive fabrication process are presented, along with challenges and future improvements. A theoretical model of the calibration potential is presented and supported with experimental results, including stability of the micro-fabricated torsional resonator and calibration demonstration on a macro-scale analog.

Chapter 6 presents the challenges of device packaging. This includes the importance
of low out-gassing materials and stress imparted from thermal mismatches between the silicon device and packaging material. The influence of packaging stress is experimentally verified through the use of electrical interrogation of the device’s resonance frequency, before and after packaging. This relationship is confirmed through finite element modeling and correlated to die attachment area through x-ray imaging of the final attachment dimensions.

Chapter 7 provides a final summary of this dissertation, along with a list of the primary contributions. A few of the many continuations of this topic are also listed.

Appendices are also supplied to support the main text, as well as aid future researchers to continue this work. Appendix A supplies a theoretical overview of inertial sensor design and control. Appendix B provides a similar overview of fabrication technologies, as well as detailed information on the silicon-on-insulator fabrication process. Appendix C supplies a series of protocols which take a user from initial device conception to obtaining gyroscope performance metrics. And finally, Appendix D contains a list of vendors that were used in this research, along with contact information and services rendered.
Chapter 2

Mode Ordering of Multi-Mass Resonators

2.1 Introduction

In this chapter, the influence of parasitic modes of resonance are discussed for anti-phase resonators in terms of common-mode rejection, energy loss, and linear range expansion. Design strategies are also compared with respect to high-frequency, distributed mass resonators and low-frequency, lumped parameter resonators. For lumped parameter devices, two types of leveraging structures are presented for tailored mode placement of the primary vibratory modes, which is supported through analytical models.

For improved common-mode rejection, it is shown that both a high in-phase resonance frequency, as well as a large frequency separation between the in- and anti-phase modal resonances is beneficial. A high in-phase resonance directly resists any common-mode input, while a large frequency separation reduces the influence of mode-coupling through fabrication imperfections. Furthermore, an experimental approach is described for quantifying fabrication imperfections of mechanically-coupled multi-mass resonators. These results are then experimentally validated by measuring the g-sensitivity of two
competing designs.

These results are of particular interest to inertial sensor design, which utilizes a lumped parameter design approach. As shown in Appendix A, these types of devices benefit from a low operational frequency, which enables complex resonator design. When a high frequency is desirable, such as when designing high-Q resonators for timing or filtering applications, this design approach becomes less desirable, resulting in the use of distributed mass resonators and beams.

2.2 Resonator Approach

A double-axis resonator, or a mass that can be deflected along two orthogonal modes of resonance, is the mechanical foundation of Coriolis vibratory gyroscopes, as shown in Appendix A. There are two main design approaches for the creation of such a resonator: 1) a distributed mass device, such as a ring or disk gyroscope, and 2) a lumped parameter device, such as a tuning fork structure. Both design approaches will be discussed and compared.

2.2.1 Distributed Mass

Distributed mass resonators are vibratory devices where the mass and stiffness of the structure occupy the same space. One of the drawbacks of such an approach is less design freedom, which creates challenges when designing application-specific structures. These types of devices generally operate using bulk-mode resonances on the order of MHz, [95], or bulk-mode resonances that have been selectively etched to reduce resonance frequency to below 100 kHz, [96], [97].
Due to high operational frequencies and low mechanical displacements, performance of these devices tend to be limited due to thermomechanical noise, as discussed in Appendix A. However, their strength is in common-mode rejection and structural symmetry. Due to their high frequency, they operate in ranges that significantly exceed that of environmental conditions, which are typically restrained to below 10 kHz [94], thus resulting in robust performance.

An example of the bulk-mode resonances common to these types of devices is shown in Figure 2.1, which reflects three mode shapes of a centrally supported disk resonator. While translational modes exist and are always the lowest frequency vibrational mode for this type of device (left), due to the overall stiffness of the structure, they are adequate at resisting external acceleration. The modes of operation are commonly referred to as wineglass modes (center and right). These resonances are balanced in momentum, contributing to lower energy loss and improved common-mode rejection, as will be discussed later in this chapter.
2.2.2 Lumped Parameters

Lumped parameter resonators are vibratory devices where the mass and stiffness of the structure are independently designed, consisting of a mass suspended by thin flexures. A basic schematic of this approach is shown in Figure 2.2. By separating these two parameters (mass and stiffness), vibrational amplitude can be designed separately from frequency, allowing greater design space depending on the final performance goals. Due to the comparatively large mass, frequency of these structures are typically below 100 kHz, with commercial products on the order of 10 kHz.

Single-mass lumped parameter resonators have a number of considerations. Unlike distributed mass resonators, their mode shape is not balanced in momentum, leading to both high energy loss through the substrate and high sensitivity to common-mode acceleration. Their low frequencies can also be considered both an advantage and disadvantage: low operational frequencies directly improve the mechanical scale factor of Coriolis vibratory gyroscopes; however, the device is more sensitive to external stimuli.

Fortunately, due to the improved design space granted by tuning-fork structures, each of these challenges can be addressed, as will be shown below.
2.3 Motivation

In the past, a significant amount of research has been conducted to optimize resonators for timing applications. High stability clocks are useful for a wide range of applications, including communications, microprocessors, and digital sampling. For these applications, both a high frequency and high Q-factor are beneficial to performance. A high frequency leads to larger potential sampling rates, while a high Q-factor increases frequency stability and reduces noise in the output signal. For these reasons, the product of a device’s frequency and Q-factor, $f \cdot Q$, has become an important figure of merit for these types of devices.

For resonators used as inertial sensors, this is not the case. While a high Q-factor is still beneficial to performance in terms of noise and stability, the optimal frequency of such a device is a bit more obscure. As shown in Appendix A, the mechanical scale factor of Coriolis vibratory gyroscopes, as well as amplitude modulated accelerometers, is inversely proportional to frequency. Furthermore, for the frequency range of interest, Q-factor can also be maximized by reducing the natural resonance frequency and thereby thermoelastic damping, as will be later discussed. However, minimizing frequency is not the answer either: a low operational frequency increases the influence of external acceleration, which is a disadvantage for gyroscopes, while also limiting the bandwidth of accelerometers. In fact, errors due to external vibration is typically ignored when characterizing commercial MEMS gyroscopes [93].

One technique to help reduce the influence of external acceleration without raising the natural frequency of the device is through the use of the anti-phase vibratory mode.
of tuning fork resonators. As will be discussed in Section 2.4, anti-phase mode shapes are present in resonators of two more more masses, each driven with equal magnitude but opposite direction. Such a resonance cancels common-mode displacement; however, when conventional flexures are used, the in-phase vibratory mode is always lower in frequency than the anti-phase mode. This leads to a higher than desired anti-phase operational mode, which reduces the mechanical scale factor, while also leaving the frequency of the in-phase vibratory mode, which is most sensitive to external acceleration, low and most likely to experience this effect.

For this reason, mode ordering can be a useful approach to improving the performance of these types of sensors. Through the use of the enhanced design space of lumped parameter resonators, the in-phase and anti-phase resonance of a tuning fork resonator can be switched in the frequency spectrum. This leads to a low anti-phase resonance with a high mechanical scale factor, while pushing the acceleration-sensitive in-phase resonance to high frequencies which the device is less likely to experience.

The rejection of external acceleration, also known as common-mode rejection, is one of three primary motivational factors in this work for pursuing enhanced mode ordering techniques. Reducing energy loss due to structural asymmetry, as well as expanding the linear range of frequency modulated accelerometers were also examined. All three factors are outlined below.

2.3.1 **Common-mode Rejection**

The anti-phase vibratory mode of a tuning fork resonator creates a first order rejection of common-mode acceleration. As previously mentioned, a low anti-phase resonance
frequency is ideal for maximizing the mechanical scale factor and Q-factor of devices within the examined frequency spectrum; however, a high in-phase resonance frequency is ideal for rejecting external acceleration.

Using the techniques discussed in this chapter, the influence of external acceleration can be minimized using four different methods:

1) **Designing the in-phase vibratory mode to be higher in frequency than the ambient environmental conditions.** Depending on the application, the power of ambient vibrations is typically maximized at frequencies below 200 Hz. This continues to taper off at higher frequencies, typically achieving a relatively full attenuation at frequencies between 5 and 10 kHz, [94]. Because these vibrations can directly actuate the in-phase vibratory mode, designing this resonance beyond the frequency spectrum of these ambient conditions reduces this energy transfer. This concept is discussed further in Appendix A, Figure A.9.

2) **Designing the in-phase vibratory mode to be as high as possible to reduce the influence of static displacement and low frequency environmental vibrations.** The in-phase vibratory mode of a tuning fork resonator behaves similar to the single-mass resonator of Figure 2.2. Any external acceleration creates a static deflection of the masses according to Equation 2.1. This displacement can be minimized by maximizing the in-phase resonance frequency.

\[
F = ma = kx
\]
\[
x_{in} = \frac{a}{\omega_{in}^2}
\]  
(2.1)
where $\omega_{in}$ is the in-phase resonance frequency, $x_{in}$ is the in-phase displacement, and $a$ is external acceleration.

Under ideal conditions, this in-phase displacement should be canceled by the anti-phase detection scheme; however, this is only a first order rejection. Due to nonlinearity of the capacitive electrodes, this in-phase displacement of the resonator can still influence acceleration sensitivity. Additional details of this capacitive nonlinearity are discussed in Appendix A.

3) Designing the in-phase vibratory mode to be as high as possible to reduce the Q-factor of the in-phase resonance. For the frequency ranges discussed in this work, thermoelastic damping is proportional to both resonance frequency and beam thickness. Both of these factors can contribute to reducing the Q-factor of this resonance, which reduces the in-phase magnitude of the device when given an external impulse. This concept is further discussed in Chapter 3, as well as Appendix A using Figure A.9.

4) Designing the in-phase vibratory mode to be as high as possible to reduce the time constant of the in-phase resonance. The time constant of a resonator is inversely proportional to the rate of decay of an impulse, and proportional to the Q-factor of the vibratory mode divided by the frequency, $\tau \propto \frac{Q}{\omega}$. Q-factor can be minimized by raising frequency, but in addition to this, raising frequency also directly reduces the damping constant of the in-phase vibratory mode. Minimizing the in-phase damping constant also minimizes the influence of sense-mode self-oscillations when exposed to an external shock for a system using open-loop control architecture.
2.3.2 Energy Loss through Structural Asymmetry

Mode ordering can not only be used to reduce the transfer of energy from the environment to the resonator, but also from the resonator to the environment, leading to a reduction in Q-factor. These two concepts are directly related, and are only separated by the direction of energy flow. As such, minimizing this energy transfer is beneficial for both cases, and as discussed, one of the most common methods of achieving this is through the use of a tuning fork structure.

Ideally, the symmetry of tuning fork structures allows the complete decoupling of the anti-phase resonance from any in-phase actuation. The challenge of such a feat arises when attempting to actually fabricate these structures. Because no fabrication process is perfect, there is always some degree of structural error across the mirrored plane of symmetry. While this error may be minimized as a fabrication process becomes increasingly refined, complete elimination is an implausible goal. Due to the narrow beam elements, stiffness mismatch dominates this source of error and when included into the system dynamics, Equation 2.2 is formed.

\[
\begin{align*}
\ddot{x}_{an} + \omega_{an}^2 x_{an} &= \frac{\Delta k}{2m} x_{in} + \frac{F(t)}{2m} \\
\ddot{x}_{in} + \omega_{in}^2 x_{in} &= \frac{\Delta k}{2m} x_{an} + a(t)
\end{align*}
\] (2.2)

where \(x_{an}\), \(x_{in}\) and \(\omega_{an}\), \(\omega_{in}\) are the modal displacement and natural frequencies of the anti-phase and in-phase modes, respectively, \(m\) is the mass of each proof mass, \(\Delta k = k_1 - k_2\) is the stiffness mismatch, \(F(t)\) is an anti-phase forcing function, and \(a(t)\) is external acceleration. For simplicity, damping is ignored in this equation for the time
being.

By ignoring the external forcing conditions and solving Equation 2.2 for the degree of in-phase / anti-phase mode-mixing for an arbitrary amplitude, Equation 2.3 is formed.

\[
\frac{x_{an}}{x_{in}} = \frac{\Delta k}{2m(\omega_{an}^2 - \omega_{in}^2)}
\] (2.3)

In this equation, it is shown that the energy coupling between the two modes of interest is a function of both the degree of stiffness asymmetry, as well as the frequency separation between the in-phase and anti-phase modes of resonance. By expanding this frequency separation, the degree of energy transfer can be minimized.

It should also be noted that for a given external acceleration, a high in-phase resonance, \(\omega_{in}\), directly reduces the invoked in-phase amplitude of the structure for a fixed acceleration, thereby also reducing the amount of energy available for transfer to the anti-phase mode. In this way, maintaining a low frequency anti-phase resonance while maximizing the in-phase frequency is the optimal strategy.

2.3.3 Linear Range Expansion

The concept of mode ordering is also applicable to tuning fork frequency modulated accelerometers, which will be further discussed in Chapter 4. In short, when choosing the resonance frequencies of a tuning fork FM accelerometer, there is a trade-off between the mechanical scale factor of the device and the linear range. A low frequency in-phase resonance creates a high resonator deflection for a given acceleration. While this high
deflection can cause a large frequency shift due to an electrostatic spring, it also causes
the resonator to snap to the electrostatic springs sooner than need be.

This trait can be avoided through mode ordering. By selectively matching the in-
phase and anti-phase resonances, the anti-phase motion of the structure becomes a part
of the snapping condition, making it difficult for both tines to snap at once. Not only
does this expand the linear range of the device, but also the mechanical scale factor,
though at the cost of nonlinearity. This concept will be discussed in much greater detail
in Chapter 4.

2.4 Anti-Phase Resonator Designs

Two types of anti-phase resonators will be discussed in this work: a single-axis device
with two degrees of freedom, and a double-axis device with eight degrees of freedom. The
single-axis tuning fork structure is a classic Coriolis vibratory gyroscope architecture that
has been present since the first MEMS vibratory gyroscopes. The double-axis tuning
fork, however, is a more recent design which applies the same design principals to both
degenerate modes of resonance, creating an even higher degree of symmetry.

2.4.1 Single-Axis, Two Degrees of Freedom

A schematic of the classic single-axis tuning fork architecture is shown in Figure 2.3. As
it can be seen in this image, the structure consists of two identical proof masses, shown
in blue, which are both grounded to the substrate with springs of equal stiffness, $k_A$. The
two proof masses are also mechanically coupled together using a spring of value $k_C$. The
structure is symmetric, with a single axis of symmetry passing through the middle, and
perpendicular to, the coupling spring, $k_C$. Each proof mass contributes a single state variable to the system in the form of displacement, which can also be transformed into relative motion as the in-phase and anti-phase vibratory modes, as depicted.

This relative motion is of greater importance than the displacement of each individual proof mass when considering the frequencies of vibration. Due to the symmetry of the structure, the coupling spring, $k_C$, does not deflect when in-phase motion is being considered; the spring simply translates as the relative motion of the proof masses remains constant. This leads to the coupling spring only contributing stiffness to the anti-phase vibratory mode. The in-phase and anti-phase resonance frequencies of such a structure are provided in Equation 2.4.

$$
\omega_{\text{In-Phase}} = \sqrt{\frac{k_A}{m}} \\
\omega_{\text{Anti-Phase}} = \sqrt{\frac{k_A + 2k_C}{m}}
$$

(2.4)

It should be noted that for conventional spring coupling, stiffness is always positive. This leads to anti-phase vibratory modes that are always higher in frequency than their
in-phase counterparts.

2.4.2 Double-Axis, Eight Degrees of Freedom

When applying a similar tuning fork design strategy to a double-axis structure, the result can be seen in Figure 2.4. This structure can be thought of as two single-axis tuning forks placed parallel to one another. By allowing each of the grounding and coupling springs to be of equivalent stiffness, along with identical proof mass dimensions, a structure can be created with symmetric properties along both axes. When examining a single axis of such a device, their are four primary modes shapes that can be seen. In order to preserve consistent nomenclature, these mode shape are defined as: in-phase, anti-phase, opposing in-phase, and double anti-phase. These shapes are present along both axes, creating a total of eight degrees of freedom.

Table 2.1 displays the relative motion of each of the four masses for each given mode shape. It should also be noted that all of these mode shapes do not couple to themselves between axes through the Coriolis force. The anti-phase and in-phase modes are symmetric in this regard; however, the opposing in-phase and double anti-phase mode shapes switch from one axis to the other when used as a vibratory gyroscope. While this is not a concern when the anti-phase mode shape is used for device operation, it does supply evidence for why the other modes should not be used. In addition, while the in-phase mode is sensitive to acceleration, it can be seen that the opposing in-phase mode is particularity sensitive to torque.

In a similar way as Equation 2.4, the resonance frequencies of each vibratory mode shapes can be calculated. The solution is provided in Equation 2.5.
Figure 2.4: Schematic of a double-axis tuning fork.

<table>
<thead>
<tr>
<th>Mode Shape</th>
<th>Mass 1</th>
<th>Mass 2</th>
<th>Mass 3</th>
<th>Mass 4</th>
<th>Coriolis Coupling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anti-Phase</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>Anti-Phase</td>
</tr>
<tr>
<td>In-Phase</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>In-Phase</td>
</tr>
<tr>
<td>Opposing In-Phase</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>Double Anti-Phase</td>
</tr>
<tr>
<td>Double Anti-Phase</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>Opposing In-Phase</td>
</tr>
</tbody>
</table>

Table 2.1: Double-axis tuning fork mode shapes and properties.
\[ \omega_{\text{In-Phase}} = \sqrt{\frac{k_A}{m}} \]

\[ \omega_{\text{Opposing In-Phase}} = \sqrt{\frac{k_A}{m}} \]

\[ \omega_{\text{Anti-Phase}} = \sqrt{\frac{k_A + 2k_C}{m}} \]

\[ \omega_{\text{Double Anti-Phase}} = \sqrt{\frac{k_A + 2k_C}{m}} \]

(2.5)

The same challenge of the single-axis tuning fork resonances exists for the double-axis tuning fork: a higher anti-phase frequency exists due to the additional stiffness of the coupling spring. However, in addition to this challenge, the double-axis tuning fork contains degenerate mode shapes along each individual axis (i.e. \( \omega_{\text{In-Phase}} = \omega_{\text{Opposing In-Phase}} \) and \( \omega_{\text{Anti-Phase}} = \omega_{\text{Double Anti-Phase}} \)). This effect leads to high energy coupling between each degenerate mode under static conditions, as well as an increased sensitivity to torque confounding the measurement.

### 2.5 Multi-Mass Coupling Structures for Tailored Mode Ordering

In order to solve these mode ordering challenges, three techniques can be used: 1) Internal spring coupling, 2) External lever coupling, and 3) Internal lever coupling. The first two techniques will be presented below.

#### 2.5.1 Internal Spring Coupling

The most common design element for vibratory MEMS are simple beams with various supporting conditions, introduced to allow deflection between structures. As shown in
Equation A.6 of Appendix A, when beams are directly used as stiffness elements, the stiffness they impart on the structure must be greater than zero. As discussed, the consequence of this condition for the double-axis tuning fork are in-phase vibratory modes with frequencies that are always less than their anti-phase counterparts, as shown in Equation 2.6.

\[
\{ k \in R : k > 0 \} \quad (\omega_{\text{Anti-Phase}} = \omega_{\text{Double Anti-Phase}}) > (\omega_{\text{In-Phase}} = \omega_{\text{Opposing In-Phase}})
\]

If the desired frequency relations for the given application can be satisfied by Equation 2.6, then no further modifications must be made. However, as previously shown, this is not conducive to vibratory gyroscopes. For gyroscopes, the anti-phase vibratory mode should be the lowest frequency mode shape, with the in-phase vibratory mode being the highest frequency. The relations of the double anti-phase and opposing in-phase are less critical, but should be each isolated in the frequency spectrum to reduce the transfer of energy due to fabrication imperfections.

2.5.2 External Lever Coupling

In order to isolate each individual mode shape within the frequency spectrum, external levers can be used. These levers can be seen as thick beam elements which are centrally anchored and externally attached to each neighboring proof mass pair. Each attachment is designed to be robust to translation, yet weak to rotation. In this way, the thickness of the lever behaves as an anti-phase stiffness element, forcing the anti-phase motion of neighboring proof masses using the stiffness of the lever. When including this external
lever coupling, the double-axis tuning fork from Figure 2.4 is shown in Figure 2.5, with lever stiffness \( k_L \).

The equations of motion can then be updated to include this additional lever stiffness, with modal frequencies calculated as shown in Equation 2.7.

\[
\begin{align*}
\omega_{\text{In-Phase}} &= \sqrt{\frac{k_A + 2k_L}{m}} \\
\omega_{\text{Opposing In-Phase}} &= \sqrt{\frac{k_A}{m}} \\
\omega_{\text{Anti-Phase}} &= \sqrt{\frac{k_A + 2k_C}{m}} \\
\omega_{\text{Double Anti-Phase}} &= \sqrt{\frac{k_A + 2k_C + 2k_L}{m}} 
\end{align*}
\]  

Equation 2.7 shows that with the inclusion of this additional design element, the four
primary modal frequencies can now be isolated from one another, which solves one of the challenges for using such a device as a vibratory gyroscope. The placement of these frequencies is still not ideal, though. Using this method, $\omega_{\text{Opposing In-Phase}}$ is now the lowest frequency vibratory mode, which is still sensitive to external torque.

In order to achieve a better representation of this frequency placement, the equations of motion can be transformed to state space of the four proof masses. Equation 2.8 shows the classic state space relationship, along with Equation 2.9 showing the transformation of the individual modal Q-factors to the damping coefficients applied to each proof mass. The complete state space parameters are then provided in Equation 2.10.

\[
\dot{x} = Ax + Bu \\
y = Cx + Du
\] (2.8)

\[
T = \begin{bmatrix}
-\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \\
\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\
\frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\
-\frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & \frac{1}{4}
\end{bmatrix}
\]

\[
\bar{C} = \begin{bmatrix}
\omega_{\text{AP}}/Q_{\text{AP}} & 0 & 0 & 0 \\
0 & \omega_{\text{IP}}/Q_{\text{IP}} & 0 & 0 \\
0 & 0 & \omega_{\text{OIP}}/Q_{\text{OIP}} & 0 \\
0 & 0 & 0 & \omega_{\text{DAP}}/Q_{\text{DAP}}
\end{bmatrix}
\] (2.9)

\[
C = T^{-1}\bar{C}T = \begin{bmatrix}
c_1 & c_{12} & c_{13} & c_{14} \\
c_{12} & c_2 & c_{23} & c_{24} \\
c_{13} & c_{23} & c_3 & c_{34} \\
c_{14} & c_{24} & c_{34} & c_4
\end{bmatrix}
\]
\[ A = \frac{1}{m} \begin{bmatrix}
0 & m & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
(-k_A - k_C - k_L) & -c_1 & k_C & -c_{12} & -k_L & -c_{13} & 0 & -c_{14} \\
0 & 0 & 0 & m & 0 & 0 & 0 & 0 & 0 \\
k_C & -c_{12} & (-k_A - k_C - k_L) & -c_2 & 0 & -c_{23} & -k_L & -c_{24} \\
0 & 0 & 0 & 0 & 0 & m & 0 & 0 & 0 \\
-k_L & -c_{13} & 0 & -c_{23} & (-k_A - k_C - k_L) & -c_3 & k_C & -c_{34} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & m & 0 \\
0 & -c_{14} & -k_L & -c_{24} & k_C & -c_{34} & (-k_A - k_C - k_L) & -c_4 \\
\end{bmatrix} \]

\[ B = \begin{bmatrix}
0 \\
F_1/m \\
0 \\
F_2/m \\
0 \\
F_3/m \\
0 \\
F_4/m \\
\end{bmatrix} \]

\[ C = \begin{bmatrix}
0 & 1 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix} \]

\[ D = \begin{bmatrix}
0 \\
\end{bmatrix} \]

(2.10)
Equation 2.10 can be used to simulate the actuation and detection of a single proof mass of the double-axis tuning fork with external lever coupling, along with the identification of each modal resonance peak. The frequency spectrum of such an experiment can be shown in Figure 2.6 for a typical design.

2.6 Internal Lever Coupling

Even though the external lever coupling is able to solve one of the challenges of the double-axis tuning fork being used as a vibratory gyroscope, the placement of the modal frequencies is still not ideal. In this section, a new internal coupling mechanism is presented which is able to replace the traditional beam spring. In doing so, the equivalent of a negative stiffness can be applied between two proof masses, greatly opening the available design space.
2.6.1 Concept

As previously discussed, the design goals for the use of the double-axis tuning fork as a vibratory gyroscope have three components: 1) A low frequency anti-phase resonance, 2) A high frequency in-phase resonance, and 3) Excellent frequency separation of the additional mode shapes. The third goal has already been achieved through the use of external levers; however, what design changes can be made to promote the first two goals?

When examining Equation 2.7 and assuming a positive stiffness for each value of $k$, it can be shown that the frequency of the opposing in-phase mode shape will always be the lowest of the four, the double anti-phase mode shape will always be the highest, and the in-phase and anti-phase mode shapes can lay anywhere between the two as a function of $k_C$ and $k_L$. However, if $k_C$ is allowed to be negative, as shown in Equation 2.11, this mode ordering changes. Assuming that $k_C \leq 0$, the frequency of the anti-phase mode shape will always be the lowest of the four, the in-phase mode shape will always be the highest, and the opposing in-phase and double anti-phase mode shapes can lay anywhere between the two as a function of $k_C$ and $k_L$. Such a situation enables all three of the original design goals.

$$\left\{ k_C \in R : k_C \leq 0 \right\} \quad (2.11)$$

This principal can be modeled through the use of Equation 2.10 by allowing a negative stiffness value to exist for $k_C$. Much like Figure 2.6, the result of this effect can be seen
This modeling shows that a negative coupling stiffness is one method to enable the ideal mode ordering of the double-axis tuning fork, for use as a Coriolis vibratory gyroscope. However, the question still remains on how a negative stiffness coupling element can be created. As will be shown in the next section, this feat can be accomplished by redirecting the stress produced by the deflection of the proof masses, through the use of levering structures.

2.6.2 Analytical Model

When placing a traditional spring between two resonant proof masses, the mechanical resistance of that spring is dependant upon only one state variable: the difference in deflection between the two proof masses. Because this single state variable is only related
to the anti-phase motion of the structure, the resistance of the spring is only applied to
the anti-phase motion, leading to an in-phase motion that is more compliant. In order
to create a negative coupling stiffness, an additional in-phase state variable must be
included, thus requiring the anchoring of the spring to the substrate. However, in order
to purposefully create a negative coupling stiffness, this anchoring must be performed in
a way which disproportionally stiffens the in-phase motion of the structure, with respect
to the anti-phase motion.

One way to accomplish this is by redirecting the motion of the two proof masses,
or primary resonator, and selectively applying it to a secondary resonator. Depending
on how this motion is transmitted to the secondary resonator, the in-phase and anti-
phase motion of the primary resonator can be decoupled by inducing two different mode
shapes within the secondary resonator. For this case, a simple clamped-clamped beam
is being used as the secondary resonator, with the first and second mode shapes of this
resonator being induced by the anti-phase and in-phase motion of the primary structure,
respectively.

A schematic of this new coupling structure is shown in Figure 2.8, along with a
model of the equivalent parameters. Due to the anchoring of the structure, an additional
grounding stiffness is applied to each proof mass which increases the total individual
stiffness applied to each mass: \( k_A = k + k_{IN} \).

Figure 2.9 shows the critical dimensions of interest of one-half of the coupling struc-
ture, which is mirrored to form the second half. The coupling structure is formed of two
distinct parts: 1) The levers with dimensions \( w \) and \( h \), which are used to transform the
Figure 2.8: Schematic of the lever coupling (top), and simplified mass/spring system (bottom).

Figure 2.9: One-half of the coupling structure with variables defined.
displacement of the proof masses, and 2) A clamped-clamped beam which absorbs this transformed displacement. Because the motion of each individual proof mass is applied to the clamped-clamped beam separately, the in-phase and anti-phase resonances of the primary structure force two different mode shapes within the secondary resonator. These modes shapes are shown in Figures 2.10 and 2.11. For anti-phase motion of the proof masses, the clamped-clamped beam is forced in the same direction by the proof masses, but torqued oppositely, Figure 2.10. For in-phase motion, the clamped-clamped beam is forced in opposite directions by the proof masses, but torqued in the same direction, Figure 2.11. Because an anti-phase resonance of the primary structure forces the first mode shape of the beam, while an in-phase resonance forces the second mode shape, it can intuitively be seen that the in-phase resonance will have a greater stiffness.
In order to confirm and quantify this variability in stiffness, analytical expressions for the anti-phase and in-phase stiffness of the coupling structure were calculated, with respect to the motion of the primary resonator. The results are shown in Equations 2.12 and 2.13 for the anti-phase and in-phase mode shapes, respectively.

\[
k_{\text{Anti-Phase}} = \left[ \frac{(L - g)^2 h}{EI(2L)^3 w} \left[ \frac{wh}{w^2 + h^2} \frac{2}{3} L^2 (L - g)(L + 3g) + h(4gL^2) \right] \right]^{-1} \tag{2.12}
\]

\[
k_{\text{In-Phase}} = \left[ \frac{(L - g)^2 h}{EI(2L)^3 w} \left[ \frac{wh}{w^2 + h^2} \frac{2}{3} g^2 (L - g)(3L + g) + h(2gL^2 - 4Lg^2 - 2g^3) \right] \right]^{-1} \tag{2.13}
\]

These equations assume that the levers completely transform the stress between the primary and secondary resonators, with zero loss. Concerning the application of this assumption, there are two components: 1) The levers are perfectly ridged, and 2) The hinge locations of the levers only transmit deflection and not torque. These assumptions can be validated for levers with large thicknesses, as well as hinge locations with narrow thicknesses. Both assumptions were verified using Finite Element Modeling and experimental data.

In order to compute the equivalent parameters of \(k_C\) and \(k_{IN}\), as defined in Figure 2.8, the ratio of the anti-phase to the in-phase stiffness can be calculated, as shown in Equation 2.14.
\[
\frac{k_{\text{Anti-Phase}}}{k_{\text{In-Phase}}} = \frac{w h}{w^2 + h^2} \frac{2}{3} g^2 (L - g)(3L + g) + h(2gL^2 - 4Lg^2 - 2g^3) \\
\frac{w h}{w^2 + h^2} \frac{2}{3} L^2 (L - g)(L + 3g) + h(4gL^2) 
\]

(2.14)

To determine the equivalent coupling stiffness, \(k_C\), this stiffness ratio can be compared to that of the single-axis tuning fork structure provided in Equation 2.4. Rearranging these equations leads to a formulation of the coupling stiffness provided in Equation 2.15. The additional anchoring stiffness, \(k_{IN}\), can be calculated in a similar way, revealing that \(k_{IN} = k_{\text{In-Phase}}\).

\[
k_C = \left( \frac{k_{\text{Anti-Phase}}}{k_{\text{In-Phase}}} - 1 \right) \frac{1}{2} k 
\]

(2.15)

When \(\frac{k_{\text{Anti-Phase}}}{k_{\text{In-Phase}}} < 1\), \(k_C\) is negative, which effectively forces the in-phase motion of the proof masses to be stiffer than the anti-phase motion. By plotting this value versus \(L/g\) for various values of \(g\), Figure 2.12, it can be seen that for \(L/g = 1\), \(k_C = -k\). As \(L/g\) increases, \(k_C\) quickly becomes less negative as it approaches \(k_C = -0.25 \cdot k\). Then, as a function of \(h, w,\) and \(g\), \(k_C\) starts becoming more negative again, approaching a value of \(k_C = -0.5 \cdot k\).

While minimizing \(k_C\) is ideal for stiffening the in-phase motion of the structure, the anti-phase motion of the structure must remain compliant to allow for high amplitudes of motion. When examining the additional stiffness the coupling structure places on each proof mass as a function of \(L/g\), Figure 2.13, it can be seen that for low values
Figure 2.12: Plot of the value of the equivalent coupling spring \((k_C)\) versus \(L/g\) for \(h = 1000\mu m\) and \(w = 100\mu m\).

Figure 2.13: Plot of the value of the equivalent coupling spring \((k_c)\) and the in-phase stiffness \((k_{IN})\) of the coupling structure versus \(L/g\) for \(h = 1000\mu m\), \(w = 100\mu m\), and \(g = 10\mu m\).
of \( L_g \), the stiffness \( k_{lN} \) increases exponentially, which will effectively increase the anti-phase frequency as well. For high values of \( L_g \), this requires the connection points on the clamped-clamped beam to be as close together as possible. This quantity is limited by the minimum feature width of the fabrication process, as well as the maximum length of the clamped-clamped beam. While it is possible to merge the connection points, this results in the motion of the proof masses not necessarily being required to transmit through the clamped-clamped beam, resulting in a break down of the analytical expressions and assumptions, deteriorating the frequency separation potential. Maximizing \( L_g \) has an additional consequence of creating a stress concentration, which can both increase the chance of device fracture and greatly increase thermoelastic damping.

### 2.6.3 Experimental Results

While evidence of this negative stiffness effect has already been provided analytically, additional results confirming this influence can be seen throughout this work.

In the following section, Figure 2.15 compares the analytical model of Equation 2.10 with experimental results of a quadruple mass gyroscope, which utilizes the internal lever coupling. These two frequency spectrums are shown to be in good agreement.

Finite element modeling is also used to support this effect, and is provided in Figures 3.11 through 3.14. These images also demonstrate the improved frequency separation between the four primary modes of resonance.

Additional results are also provided in Figures 3.24 through 3.28, which experimentally identify the individual mode shapes within the multi-peak frequency spectrum. This is performed by driving the structure from a single mass and detecting the motion from
all four masses. By taking multiple sweeps and modifying the detection polarity applied
to each mass, the mode shape of each frequency peak could be identified.

2.7 Fabrication-Induced Asymmetry

As previously mentioned, fabrication-induced asymmetry within the structure of the
resonator can lead to an increase in the energy transfer between each of the modes of
resonance. This effect was shown for a single-axis tuning fork in Equation 2.3, where
the stiffness asymmetry of the individual anchoring springs of each mass, $\Delta k$, is directly
related to the amplitude coupling of the in-phase and anti-phase vibratory modes. A
similar effect is true for the double-axis tuning fork, though with increased complexity
due to the additional mode shapes.

So far, this energy transfer has been minimized passively by increasing the frequency
separation of each of the modes of resonance, as well as raising the frequency of the
mode shapes which are sensitive to common-mode acceleration and torque. However, if
active compensation is wished to be employed, the degree of stiffness asymmetry must be
experimentally determined. This poses a challenge due to the multiple discrete elements
of the double-axis tuning fork structure, all of which are mechanically coupled, which
makes the identification of each individual parameter difficult.

To meet this need, the analytical model of the resonator can be used, as shown in
Equation 2.10. By selectively introducing fabrication asymmetries within the model,
the effects can be observed in the simulated frequency sweeps of the device. These
influences can then be linearized and compared to the experimental data to determine
the stiffness variations across the device. Once complete, this information can then be
used to selectively tune the individual proof masses using the electrostatic stiffness effect, shown in Equation A.17. This method directly reduces the stiffness asymmetry through active compensation, thus reducing the energy transfer between the modes of resonance.

2.7.1 Mode-Mixing

The deflections of each individual mass and the deflections of the four primary mode shapes of the structure are analogous methods of representing the same information. As such, a transformation exists between these two sets of variables. This transformation was used previously to convert the Q-factors of the individual mode shapes to the damping coefficients of the model, as shown in Equation 2.9. In a similar way, this same transformation can be used to convert between these two sets of state variables, as shown in Equation 2.16.

In Equation 2.16, the transformation $T$ represents a system with ideal symmetry, and therefore ideal modal decoupling. While the definitions of these modal resonances always reflect this idealized behavior, the truth is that the real transformation matrix has slight variations as a function of the structural asymmetry. In this way, the real modal resonances still span the four dimensional space; however, this transformation matrix may be slightly tiled by an arbitrary angle around any of the four principal modal axes.

A similar phenomena is discussed in Appendix A for the degenerate axes of a single mass gyroscope, which is eventually represented in Equation A.51. This is an example of the same principal, but perhaps easier to understand because there is only one modal resonance, compared to the four resonances of Equation 2.16.
There are four unknown variables when considering modal asymmetry, and these variables can be considered in one of two ways: 1) The rotation of the transformation matrix around each principal modal axis, and 2) The stiffness variation from the nominal position of each individual mass. The second method assumes that any influence of mass asymmetry is small compared to stiffness, which is a valid assumption for these types of structures. In order to represent this information as a physical quantity, which can also be used for eventual frequency tuning of the stiffness asymmetry, the second representation was chosen.

A schematic model of these four stiffness variations are shown in Figure 2.14 as $\Delta k_1$, $\Delta k_2$, $\Delta k_3$, and $\Delta k_4$. Because these unknowns represent a variation from the nominal stiffness, $k_A$, a condition exists such that their summation must equal zero. This condition can be used to reduce the number of unknowns from four to three. A similar condition would exist if these variables were represented as rotations of the transformation matrix.
2.7.2 Asymmetry Quantification

There are only a few methods that can extract these stiffness asymmetries from experimental data, each with limited success. Perhaps the most intuitive method is to symmetrically force the structure as if under normal operation, examine the voltage response at each of the four individual proof masses, and correlate these sense voltages to displacement using the known capacitances of the electrodes. These displacements can then be compared to a model of the resonator and the individual anchoring stiffness of each proof mass determined. The challenge with this approach is that it assumes that there is zero error within the electrostatic electrodes. Because the designed capacitance is being used to extract the displacements of each individual mass, any variation within
these electrodes across the device can not be isolated and is included within the measured displacement variation, obscuring the measurement. For this reason, this solution is not ideal.

An alternative approach is to use the fact that multiple resonance peaks appear when only a single mass of the structure if forced throughout the frequency spectrum of interest. By actuating the structure from a single mass and observing the resonance peaks from each of the remaining masses, a selective ratio of the observed resonance peaks can be used to determine the stiffness variations across the die. Because a ratio of the magnitudes of the observed peaks is the output of each capacitor, it is a dimensionless number that is independent of any capacitive variations across the die. An example of how well the analytical model provided in Equation 2.10 fits an experimental sweep of a fabricated device is provided in Figure 2.15. The device used in this experiment was a quadruple mass gyroscope, fabricated using the SOI fabrication procedure, as outlined
Figure 2.16: Comparison between the ideal model (driving M1 and sensing M2) of the device against the same model with an exaggerated 20% increase in stiffness on a single mass ($\Delta k_1 = 1.2 \cdot k_A$).

Knowing from Figure 2.15 that the model of the device accurately represents the motion of the structure, modifications to the model can be made to observe how the motion of the structure would change. By selectively tuning the anchoring stiffness of each proof mass separately, it can be observed that the magnitudes of the resonance peaks change. In addition to this, as the stiffness of each proof mass is altered, some peaks rise, while others decrease. An example of this can be seen in Figure 2.16 for a double-axis tuning fork being driven from proof mass 1 and sensed from proof mass 2. The blue curve is the ideal behavior, while the green curve is the behavior when there is a 20% increase in the anchoring stiffness of proof mass 1. As observed, due to the imperfection, two of the peaks rise while the other two decrease. A similar influence can be seen for various combinations of the driving mass, the sensing mass, and the location in Appendix B.
Figure 2.17: Direction of change in magnitude of each resonance (Rn) versus location of imperfection ($\Delta k_n$), actuating mass (An), and sensing mass (Sn). Minimally, three sweeps are needed to identify all four imperfections (one orange, green, and blue).

From Figure 2.17, it should be noted that there is a bit of redundancy; obtaining every combination of sweep is not necessary for extracting the unknown imperfections. Because there are only three unknowns, only three frequency sweeps are required to extract this information; however, the user must be selective on which sweeps are gathered. Figure 2.17 is color coded based on which data sets contain independent information. The user must gather at least one data set of each of the three different colors to span the unknown quantities. The simplest way to accomplish this is to actuate proof mass 1, then record the frequency responses of proof mass 2, 3, and 4 separately. The magnitudes of each of the four peaks must then be added or subtracted, based on the direction indicated in
Figure 2.18: Imperfections identified from a single device, showing 1.4% horizontal and 0.4% vertical change in stiffness across the device, with lower stiffness towards the perimeter of the wafer.

Figure 2.17, to obtain three different values. These values can then be transformed into the individual stiffness variations of each proof mass through a sensitivity analysis of the resonator model.

Please note that while it should be possible to create a analytical expression to directly convert this experimental data to the anchoring stiffness variations of each proof mass, such an expression is significantly complex concerning the high order of the dynamic model. Considering the dynamic motion of the structure, a total of eight state variable exist, making a sensitivity analysis significantly less computationally intensive.

For the device in question, this analysis was conducted, revealing a significant difference in the horizontal and vertical stiffness variations across the die. Horizontally, the stiffness of the die varied on the order of 1.5%, while vertically, the stiffness varied only 0.5%. When comparing this result to the location of the die on the wafer during etching,
as shown in Figure 2.18, the results appear to correlate with one another, supporting the conclusion that the radial distance of the die from the center of the wafer is a significant contributing factor of structural asymmetry. In this case, every die location away from the center of the wafer appears to contribute an approximately 0.5% stiffness variation across the axis of which the device is offset. This rule appears to be true for both the $x$- and $y$-axis of the structure, and should be able to be minimized by either a die size reduction, or an increase in the etching uniformity across the wafer.

### 2.8 Common-mode Rejection

As mentioned throughout this chapter, one of the primary motivational factors for gyroscope mode ordering is the rejection of common-mode acceleration. The primary purpose of this is to protect the sense mode from non-Coriolis induced acceleration, for such forces directly contribute to measurement uncertainty. Ideal mode ordering accomplishes this by both reducing the conversion of external acceleration to in-phase displacement, through the use of a high in-phase resonance frequency, as well as reducing the transfer of energy between the in-phase and anti-phase mode, by creating a high frequency separation between these two mode shapes.

In order to demonstrate this effect, two single-axis tuning fork devices were experimentally tested: one with a traditional spring coupling structure, and a second with the described lever coupling. Beyond the coupling structure design and in-phase resonance frequency, both devices were nearly identical. In order to quantify their degree of sensitivity to external acceleration, both devices were placed on a linear shaker and given a precise amplitude of vibration at 100 $Hz$ along the sense-mode of the resonator. By
examining the frequency spectrum of both devices, acceleration sensitivity was calculated [98], revealing over a 20 times reduction due to the modified design. This experimental result is consistent with the analytical equations.

2.8.1 Analytical Expression

When comparing resonator designs, it can be seen from Equations 2.1 and 2.3 that the influence of acceleration on the anti-phase resonance is a function of both $[\omega_{in}^2]^{-1}$ and $[\omega_{an}^2 - \omega_{in}^2]^{-1}$. $[\omega_{in}^2]^{-1}$ is the scale factor for converting applied common-mode acceleration to in-phase displacement, specifically for quasi-static input accelerations with respect to the in-phase resonance frequency. $[\omega_{an}^2 - \omega_{in}^2]^{-1}$ is the frequency dependant contributions for the conversion of in-phase displacement to anti-phase motion. In this way, an analytic expression can be formed to represent the change in the acceleration sensitivity of a resonator as a specific result of mode ordering, assuming all other factors remain constant. This expression is provided as Equation 2.17.

$$\frac{\bar{\Gamma}_{\text{Lever}}}{\bar{\Gamma}_{\text{Spring}}} \propto \frac{[\omega_{in}^2][\omega_{an}^2 - \omega_{in}^2]}{[\omega_{in}^2][\omega_{an}^2 - \omega_{in}^2]}$$ (2.17)

where $\bar{\Gamma}$ is the acceleration sensitivity of each resonator.

Equation 2.17 assumes that all factors other than resonance frequency remain constant between the two resonator designs. A few of these factors can be identified from Equation 2.3, which includes the mass of the resonator and spring asymmetry, but additional factors include nuances within the structural design of the resonator and how it is being electrically interfaced [99].
For this experiment, two nearly identical single-axis tuning fork resonators are used. Both resonators were fabricated and packaged as high-Q devices, as outlined in the SOI fabrication process of Appendix B. The only substantial different in the design of the two resonators was how the masses were coupled to one another. The resonator with the spring coupling was shown to have an anti-phase resonance frequency of 2.16 $kHz$, along with a in-phase frequency of 2.03 $kHz$. The resonator with the lever coupling had an anti-phase resonance frequency of 1.80 $kHz$, along with an in-phase frequency of 3.02 $kHz$. By comparing these resonators using Equation 2.17, the device with the levered coupling is estimated to experience approximately 25 times less acceleration sensitive, when compared to the spring coupled design.

### 2.8.2 G-sensitivity through Applied Vibration

While vibratory MEMS are known for poor performance in environments with high shock and vibration, attempting to quantify this sensitivity for static, low magnitude conditions, such as the classic 2 $g$ tip-over test, would require the resonator to have an extremely low level of short term bias to produce an acceptable resolution. For this reason, in order to compare the acceleration sensitivity of the two structures with a high degree of accuracy, a high magnitude external vibration was induced to the structure. This method is well known for characterizing the acceleration sensitivity of resonators used as frequency references [98]; however, is also applicable to the amplitude error at the operational frequency of the resonator.

Characterizing this error first involves oscillating the resonator at its operational frequency with a constant amplitude of motion. The resonator is then placed on a linear
shaker, with the axis of external vibration aligned to the sensitive axis of the resonator. An external vibration with a consistent magnitude and frequency is applied to the device, and the frequency spectrum of the output of the resonator is recorded. By centering the frequency spectrum upon the operational frequency of the resonator, sidebands appear within this spectrum, set at a frequency offset of plus and minus the frequency of the external acceleration. The difference in magnitude between the primary resonance peak and the sidebands can then be used to identify the acceleration sensitivity according to Equation 2.18.

\[
L(f_V) = 20 \log_{10} \left( \frac{\bar{\Gamma} \cdot \bar{a}}{2f_V} \right) \quad (2.18)
\]

where \( f_V \) and \( \bar{a} \) are the frequency and amplitude of the physical vibration, \( f \) is the frequency of the resonator, and \( \bar{\Gamma} \) is the acceleration sensitivity of the resonator.

While this procedure was designed for frequency references, it also applies to how static acceleration influences the amplitude of the operational frequency of the resonator. As the frequency of the external acceleration is reduced, the distance of the sidebands narrow in the frequency spectrum, until they overlap upon the operational frequency. Because this input is indistinguishable in the frequency domain, its influence is demodulated within the response of the resonator and contributes to error.

### 2.8.3 Experimental Results

The above procedure was implemented for the two types of resonators described: one with a spring coupling, and another with a lever coupling. External acceleration was also
applied with a magnitude of $4 \, g$ and frequency of $100 \, Hz$. A photograph of the test setup is shown in Figure 2.19 (right), along with the frequency spectrum of the resonators in Figure 2.19 (left). As can be seen in these results, the magnitude between the primary resonance and sidebands was $82 \, dB$ for the spring coupled device and $110 \, dB$ for the level coupled resonator. Using Equation 2.18, this corresponds to acceleration sensitivities of $2 \times 10^{-6} \, g^{-1}$ and $9 \times 10^{-8} \, g^{-1}$, respectively. Therefore, the lever coupling has been shown to improve acceleration sensitivity by approximately 22 times, which is within a 10% of the predicted value of Equation 2.17.
2.9 Conclusion

When designing resonant MEMS structures, there are two distinctly different design methodologies: 1) Distributed mass structures, and 2) Lumped parameter structures. There are multiple benefits and disadvantages to each choice; however, lumped parameter structures fundamentally have the potential for higher performance due to their lower operational frequencies and high amplitudes of motion. A significant disadvantage to this; however, is common-mode sensitivity.

In order to mitigate this sensitivity, the use of anti-phase resonances is a well known technique that can improve performance. However, this technique is limited by the asymmetry introduced by fabrication imperfections. While directly reducing this asymmetry through process optimization is certainly the best option, this can be a challenging goal for microfabrication processes.

Another option that can be used to passively reduce the influence of this asymmetry is through the use of mode ordering. This can be accomplished by raising the in-phase frequency as high as possible, while leaving the useful anti-phase frequency compliant. In this chapter, a novel coupling structure is introduced that is capable of accomplishing this goal, which is believed to be capable of achieving the highest in-phase / anti-phase frequency separation of a tuning fork structure to date. This structure has also been shown experimentally to be capable of reducing acceleration sensitivity by over 20 times, which is consistent with predicted results.

A second method of reducing asymmetry is through active stiffness compensation of the individual proof masses of the resonator. One of the greatest challenges of this
technique is the identification of the magnitude and location of these stiffness asymmetries for multi-mass devices, such as the double-axis tuning fork. It is shown that by actuating a single proof mass of the resonator and detecting the responses of the additional proof masses, the asymmetries can be derived by comparing the results to a model of the structure. This method functions independently of any capacitive variance across the device, and was used to identify a radial asymmetry across a silicon wafer, most likely due to device etching.
Chapter 3

Quadruple Mass Gyroscope

3.1 Introduction

As shown in Chapter 2, tailored mode ordering of anti-phase resonators can enhance common-mode rejection, reduce energy loss from structural asymmetry, as well as expand the linear range and sensitivity of certain types of accelerometers. In this chapter, the anti-phase, double-axis, eight degree of freedom resonator previously discussed will now be used as a Coriolis vibratory gyroscope for rate detection. Apart from resonance frequency, a number of other qualities are critical for optimal device performance: high quality factor of both modes of resonance, linearly isolated capacitive electrodes, and maximization of coupled mass are just a few. The design progression of the quadruple mass gyroscope is presented, along with the identification and minimization of the primary sources of damping for low frequency resonators (0.3 kHz to 30 kHz). These parameters are then experimentally verified and devices are interfaced with front-end electronics and control algorithms for rate characterization. An in-run algorithmic compensation method is also discussed for the elimination of rate demodulation phase noise, electronic thermal drift, and external acceleration.
3.2 Structural Design

The quadruple mass gyroscope is a complicated structure with many interacting design elements. A few of these elements include the coupling mechanisms discussed in Chapter 2, the tradeoffs in the shuttle design for axis decoupling, the degree of tolerance when attempting to preserve symmetry, the choice of gap dimensions while considering fabrication tolerances, as well as the choice of electrode structures and wire bonding. Each of these influences will be discussed below. For additional information on the details of the design process, see Appendix A.

3.2.1 Shuttles for Axis Isolation

As previously shown in Chapter 2, the quadruple mass gyroscope consists of four resonance proof masses for balanced actuation, with each mass able to deflect in both axes for proper operation. When attempting to include capacitive electrodes for driving and sensing the motion of the structure, a challenge arises: each electrode should be sensitive to only a single axis of motion. When the \(x\)- and \(y\)-axis electrodes experience cross-axis sensitivity, for instance due to lateral proof mass motion, it becomes difficult or impossible to fully decouple the motion. This results in questionable amplitudes of motion and high quadrature values which can degrade sensor performance due to time-varying phase delays during demodulation.

In order to reduce this form of error, additional single-axis shuttles are attached to the four sides of each proof mass. The motion of each proof mass must then transmit through these shuttles before interacting with the capacitive electrodes or the other proof masses. In addition to aiding the capacitive electrodes to achieve single-axis sensitivity, shuttles
also help lock the principal axes of stiffness to the sensitive direction of the capacitive electrodes, thus reducing the quadrature error.

The cost of these benefits include increased design complexity, as well as reduction of the angle gain by including mass which does not contribute to the Coriolis force. In order to maintain the balanced resonance and maximize the capacitive electrodes, four shuttles are included on each proof mass. For an initially four-mass system, should the shuttles be included in the model, this increases the number of resonance masses to twenty. Through proper spring design, it is assumed that each shuttle is perfectly locked to the corresponding axis of motion of the attached proof mass. This is a necessary assumption for device operation and can be supported through finite element modeling and optimal measurements, though the resolution of the latter is typically restricted to approximately 100 $\text{nm}$, leaving much to be desired. These shuttles must be designed to be rigid enough to force single-axis motion, while minimizing mass to maximize the angle gain of the sensor.

### 3.2.2 Capacitive Electrodes

There are two major types of capacitive electrodes to choose from: parallel plate and comb electrodes, each with the additional choice of single-sided or differential configuration. When comparing the two, parallel plate electrodes have a greater sensitivity per unit area they require, while comb drives have improvements of increased linearity for a given amplitude of motion. For this reason, parallel plates are typically used for sensing small amplitudes of motion, where the nonlinearity has little effect, while comb drives are used when high amplitudes of motion are present. While it has been shown that comb
drives can contribute to the bias of a gyroscope when used to drive the structure [76], sufficiently rigid shuttles can negate this effect by resisting the resultant cross-axis force, thus reducing this concern from the design.

Throughout the design progression of the quadruple mass gyroscope, the capacitive electrodes experienced the most intense modifications. The details of these changes are listed in Table 3.1, but a summary of the main concerns include proper electrode choice while maintaining device symmetry, as well as minimization of the number of necessary wire bonds.

### 3.2.3 Assumptions

Appendix A contains a thorough derivation of the equations of motion and assumptions for a single mass of the dynamic system. The final result is shown in Equation 3.1, which is a repeating of Equation A.62.

\[
\ddot{x} + \left( \frac{2}{\tau} + \Delta \left( \frac{1}{\tau} \right) \cos(2\theta) \right) \dot{x} + \left( \omega^2 + \omega \Delta \omega \cos(2\theta) \right) x = \frac{F_1 \cos(\alpha)}{m_a + m_c} + \frac{F_2 \sin(\beta)}{m_a + m_c} \\
+ \left( 2\Omega \frac{m_c}{m_a + m_c} \right) \dot{y} - \Delta \left( \frac{1}{\tau} \right) \sin(2\theta) \dot{y} - \omega \Delta \omega \sin(2\theta) y + T_x(\Delta k, \Delta c) \\
\ddot{y} + \left( \frac{2}{\tau} - \Delta \left( \frac{1}{\tau} \right) \cos(2\theta) \right) \dot{y} + \left( \omega^2 - \omega \Delta \omega \cos(2\theta) \right) y = \frac{F_2 \cos(\beta)}{m_b + m_c} - \frac{F_1 \sin(\alpha)}{m_b + m_c} \\
- \left( 2\Omega \frac{m_c}{m_b + m_c} \right) \dot{x} - \Delta \left( \frac{1}{\tau} \right) \sin(2\theta) \dot{x} - \omega \Delta \omega \sin(2\theta) x + T_y(\Delta k, \Delta c)
\] (3.1)

The assumptions of Equation 3.1 include:

1. Linear rigid body dynamics
2. Rayleigh damping
3. The principal axes of mass are aligned to local frame $O_{xyz}$ with consideration to external forces (i.e., $F_1$, $F_2$, and the Coriolis force).

4. The rigid body does not rotate ($\frac{\Delta k_x}{k_x} \geq 0.001$, or $\omega_n \geq 5 \omega_n$ for $\frac{\Delta k_x}{k_x} \approx 1.5\%$)

5. The satellite masses ($m_a$ and $m_b$) are locked to the motion of a single-axis of the primary proof mass ($m_c$).

6. Out-of-plane, z-axis motion is negligible (Fabrication aspect ratio of $\geq 10$)

7. Acceleration is negligible (Anti-phase internal dynamics)

8. $\frac{1}{2} \Omega \ll \omega_n \text{ (for } \omega_n \geq 1 kHz: \Omega \leq 2 Hz)$

9. $\frac{1}{22} \Omega \ll \omega_n \text{ (for } \omega_n \geq 1 kHz: \frac{\Omega}{\omega_n} \leq 2 Hz \cdot \Omega)$

   For a complete model of the four mass system, Equation 3.1 can be combined with Equation 2.10 from Chapter 2; however, this adds considerable complexity to the dynamic motion, quadrupling the number of state variables. For this reason, it is typically beneficial to assume that the state variables of $x$ and $y$ in Equation 3.1 are indeed the anti-phase motion of the four mass system. Implications of the four mass system on the anti-phase motion can then be separately analyzed and included, such as mode mixing effects on Q-factor and acceleration sensitivity.

3.2.4 Design Progression

Throughout this work, there have been three major design iterations of the quadruple mass gyroscope: Phase 0, Phase 1, and Phase 2. A side-by-side visual comparison of the
three designs is provided in Figure 3.1, with additional numerical parameters supplied in Table 3.1.

For the Phase 0 design, Q-factor maximization through symmetry was the primary goal. At this point, the concept of balancing linear and angular momentum for reducing substrate energy dissipation was still an unverified concept; it was unknown if such a mechanism would work, and if so, what the tolerance would be. For this reason, complete proof mass symmetry was employed. Each proof mass of the structure were identical to each other, as well as with themselves for each 90° rotation. While this may not be a challenge for the actual mass of the structure, it is a challenge when the electrodes are considered. This design required identical electrode structures for both the drive and sense axes, for fear the mass variations of different electrode structures would unbalance the structure and reduce the Q-factor. In addition, assuming that fabrication asymmetries alone would be enough to reduce the Q-factor, independent tuning electrodes were placed on either end of the shuttle, each with a large moment arm with respect to the center of mass. By independently controlling the voltage of each electrode, not only
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Phase 0</th>
<th>Phase 1</th>
<th>Phase 2</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Structure</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External Levering</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Internal Coupling</td>
<td>Spring</td>
<td>Lever</td>
<td>Lever</td>
<td></td>
</tr>
<tr>
<td>Springs (Feature / Gap)</td>
<td>10 / 10</td>
<td>10 / 10</td>
<td>10 / 7</td>
<td>µm</td>
</tr>
<tr>
<td>Parallel Plates (Feature / Gap)</td>
<td>15 / 7</td>
<td>24 / 7</td>
<td>14 / 7</td>
<td>µm</td>
</tr>
<tr>
<td>Comb Teeth (Feature / Gap)</td>
<td>5 / 5</td>
<td>5 / 5</td>
<td>N/A</td>
<td>µm</td>
</tr>
<tr>
<td><strong>Electrodes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Bondpads</td>
<td>65</td>
<td>25</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>Electrode Type (Drive)</td>
<td>Comb (Diff)</td>
<td>Comb (Sing)</td>
<td>Plate (Diff)</td>
<td></td>
</tr>
<tr>
<td>Electrode Type (Sense)</td>
<td>Plate (Sing)</td>
<td>Plate (Diff)</td>
<td>Plate (Diff)</td>
<td></td>
</tr>
<tr>
<td>Capacitive Gradient (Drive)</td>
<td>0.7140</td>
<td>0.3825</td>
<td>1.0726</td>
<td>µF/m</td>
</tr>
<tr>
<td>Capacitive Gradient (Sense)</td>
<td>1.1651</td>
<td>0.7806</td>
<td>1.0726</td>
<td>µF/m</td>
</tr>
<tr>
<td><strong>Packaging</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Package Type</td>
<td>DIP 24</td>
<td>DIP 24</td>
<td>LCC 44</td>
<td></td>
</tr>
<tr>
<td>Die Attachment Size</td>
<td>Full</td>
<td>2 mm ⌀</td>
<td>2 mm ⌀</td>
<td></td>
</tr>
<tr>
<td>Wire Bonding</td>
<td>Partial</td>
<td>Full</td>
<td>Full</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1: Summary of QMG design progression.

could the linear stiffness of each proof mass be manipulated, but also the torque. The remaining shuttle area consisted of differential comb drive electrodes for preserving a linear, high amplitude of motion.

The result of this design is symmetric device Q-factors in excess of one million. This result did not appear for every fabricated device, but was apparent for any device without serious fabrication or packaging errors; individual mass tuning was never required. Another interesting phenomena of these high-Q devices were self-oscillations; the anti-phase motion could easily be actuated from an external impulse, allowing the structure to experience observable ringing on the order of minutes. Published work seemed to indicate that despite the use of anti-phase motion, which is ideally robust to any common-mode impulse, small asymmetric errors can lead to external acceleration and impulse sensitivity.
Acceleration sensitivity can be a major challenge for high-Q gyroscopes. While closed-loop force-to-rebalance control methods are typically used to cancel such effects, this method is only relevant for reducing the amount of time this effect has on the output of the device; in no way does it prevent the initial false measurement. As discussed in Chapter 2, there are two ways to prevent this effect: 1) Directly reduce the degree of the fabrication imperfections, or 2) Increase the frequency separation between the anti-phase and parasitic vibratory modes.

For the Phase 1 design of the quadruple mass gyroscope, the proposed coupling structure of Chapter 2 was included in the design. This design also placed less emphasis on device symmetry by basing the structure off of modal symmetry rather than the symmetry of each individual structural mass. The result was a structure with two different electrode designs: internal electrodes and external electrodes. Even though each proof mass was no longer self-symmetric, the hypothesis was that the critical symmetry for reducing substrate losses was how each proof mass interacted with the other three masses. This concept will also be known as modal symmetry, which was guided by the lack of tuning required by the Phase 0 device to achieve a high Q-factor. An added benefit of this concept was that two distinct electrode designs could be created: one for driving the structure with comb drives, and a second for detecting the resulting motion with parallel plates. The elimination of the single-sided tuning plates also drastically reduced the number of wire bonds required for device operation; 65 bond pads were present in the Phase 0 design, which was reduced to 25 bond pads for Phase 1. Because these
devices were being placed into packages with only 24 leads, one of the main challenges of the Phase 0 device was that many of the electrodes were left unconnected. This is not ideal for device performance, due to the potential for capacitive charging. By connecting several bond pads to the same packaged pins, the Phase 1 design ensured that no electrodes would remain ungrounded; however, required a choice between differential configuration for the parallel plates of comb drives. Because of the low voltages required to actuate such a high Q-factor device, the parallel plates were chosen for the differential configuration to maximize the available capacitive gradient. The final modification made to the Phase 1 design was the reduction of the packaged die attachment area for thermal stress isolation, though this will be discussed in greater detail in Chapter 6.

After fabrication, many of the design hypothesises and modifications of the Phase 1 design turned out to be successful. Despite relying on modal symmetry, high Q-factors in excess of one million could still be obtained. In addition, the new coupling mechanism created what is believed to be the current record for in-phase / anti-phase frequency separation of a tuning-fork device. Structure packaging and testing was simplified by the elimination of superfluous wire bonds, and lead conformity was attained through a standardized wire bonding procedure. Finally, low-stress and void-free die attachment was also obtained.

Despite these successes, there was one complication: the long parallel plates led to an increase in stiction during release. During the design phase, it was assumed that the length to width ratio was the only figure of merit for determining parallel plate beam dimensions. This is surely an important figure for determining the resonance frequency
of the individual beams; however, it is not the only figure of merit. Stiction forces caused by evaporating liquid act along the entire length of the beam; as beams grow in length, the moment caused by this force becomes more pronounced, thus leading to stiffer beam requirements. By keeping the length to width ratio of the beams constant and increasing the overall length, the risk of stiction increases. For this reason, a vapor phase release method had to be employed, or at least a liquid release with a critical point drying step. Both methods were used for release of these devices and both were successful.

Stiction issues were solved in the final, Phase 2, version of the quadruple mass gyroscope, along with several other improvements. One of the primary changes was the inclusion of two anchoring points of each parallel plate electrode. This change allowed for the reduction of the beam lengths for robustness to stiction, while maintaining the same length to width ratio. With the reduction of the beam widths, this also allowed for a greater number of parallel plates for the given area, increasing the capacitive gradient by nearly 40%, despite the inclusion of an additional anchor. The package of the Phase 2 device was also changed from a dual in-line package (DIP) with 24 leads, to a leadless chip carrier (LCC) with 44 leads. The LCC package has a number of benefits. First, it is symmetrical along the $x$- and $y$-axis, leading to an increase in $x/y$ thermal stress symmetry of the package itself. Second, it does not include any signal routing, thus increasing the degree of symmetry of any parasitic capacitances or package resistance before interfacing the Printed Circuit Board (PCB). And finally, the number of available lead has increased from 24 to 44, thus allowing full differential electrodes, each capable of individual interrogation. Comb drives were also completely removed from the design, in
the interest of attaining identical electrodes for both driving and sensing the structure. The main motivation for this is to experiment with a new calibration technique known as mode-reversal, in which both the drive and sense electrodes are routinely switched for long-term calibration; however, this technique has yet to be fully tested. The final structural change was to the spring gaps. It is believed that footing effects during device etching is the primary cause of innate frequency mismatch of symmetric designs, such as the QMG. By reducing the gap at the spring locations to the minimal gap of the parallel plate electrodes, the time required to etch the springs will increase, thus reducing the amount of time they are over-etched and reducing the degree of lateral etching at the oxide interface. Preliminary results show that while the etching time of the springs has definitely increased, the parallel plates are still the last structure to fully etch through, thus still resulting in some degree of spring footing. This appears to be due to the fact that the parallel plate structures, despite having the same gap width at the springs, their increased length results in a slower etch rate due to a lower influence of edge effects. In order to allow the springs to be the last structure etched through, their gap width must be the smallest feature on the die.

For additional detail into the geometric design of each iteration, detailed views of the complete structure, single mass, as well as shuttles and internal coupling structure is supplied for each design: Phase 0 is reflected in Figures 3.2 through 3.4, Phase 1 is reflected in Figures 3.5 through 3.7, and Phase 2 is reflected in Figures 3.8 through 3.10.
Figure 3.2: Complete Phase 0 device.  

Figure 3.3: Single mass of Phase 0 device.  

Figure 3.4: Coupling and inner electrodes between left masses of Phase 0 device.
Figure 3.5: Complete Phase 1 device.

Figure 3.6: Single mass of Phase 1 device.

Figure 3.7: Coupling and inner electrodes between left masses of Phase 1 device.
Figure 3.8: Complete Phase 2 device.

Figure 3.9: Single mass of Phase 2 device.

Figure 3.10: Coupling and inner electrodes between left masses of Phase 2 device.
3.2.5 Modal Resonance Frequencies

Finite element modeling was conducted for each design iteration prior to fabrication. While the modifications between the Phase 1 and Phase 2 designs had little influence on resonant structure itself, the modifications between Phase 0 and Phase 1 were substantial. To complete this modeling, the device geometry was imported into COMSOL Multiphysics from L-Edit, which had been used to fabricate the photolithography mask for sensor definition. For model simplification, certain features which had little effect on the dynamic structure, yet would contribute extensively to the complexity of the mesh geometry were eliminated, such as etch holes of the proof masses and electrode structures. The material constants used for the structure are shown in Table A.2 and Equation A.1, which reflects single crystal silicon with crystal orientation consistent with standard \(<100>\) silicon wafers. The modeling results for the four primary modes of resonance for the Phase 0 and Phase 1 designs are shown in Figures 3.11 and 3.12, respectively.

The primary goal of the Phase 1 structure was to retain the low anti-phase operational frequency of approximately 2.6 \(kHz\), while raising all other parasitic vibratory modes. Finite element modeling of both structural designs shows this to be case. A degenerate resonance exists between the drive and sense axes of both the Phase 0 and Phase 1 designs with frequency of approximately 2.6 \(kHz\). In addition, due to the inclusion of the new internal coupling structure of the Phase 1 device, each of the parasitic resonances were raised in comparison: opposing in-phase, in-phase, and double anti-phase (from left to right in Figures 3.11 and 3.12).

Due to stiffness contributions of both the internal and external levering structures,
Figure 3.11: Finite element modeling of Phase 0 QMG with spring coupling.

Additional parasitic resonances exist:
- 2516 Hz
- 3110 Hz
- 3190 Hz

Figure 3.12: Finite element modeling of Phase 1 QMG with lever coupling.
the total in-phase / anti-phase frequency separation of the Phase 1 device was increased to 120%, which is believed to be the record for a tuning-fork device. This can be seen in comparison to the 25% frequency separation for the Phase 0 device, as shown in Figure 3.13, which highlights the modifications to the internal coupling structure.

However, these four resonances are simply the lowest frequency vibratory modes, and hence, the ones most likely encountered during operation. There are an infinite number of higher order modes which have the potential to interfere with device performance. For the quadruple mass gyroscope, the following modes of resonance are all torsional, as shown in Figure 3.14 for the Phase 1 device. According to Equation A.54 of Appendix A, to maintain at least three order of magnitude amplitude separation between the two degenerate modes of resonance in a fabrication process with a 1.5% asymmetry tolerance, the torsional resonances must be greater than or equal to five times greater than the
Figure 3.14: Finite element modeling of additional, higher order mode shapes of the Phase 1 QMG.

operational frequency. When compared to the modeling result of Figure 3.12, there is only a four times separation. While this is not an ideal quality for the device, any spring modification techniques to further stiffen the torsional resonances has resulted in an increase of either the theoretical thermoelastic damping or overall frequency of the structure; neither of these options present a viable alternative. As discussed in Appendix A, the consequence of low torsional vibratory modes is an increase in device quadrature and potentially bias.
3.3 Quality Factor and the Primary Mechanisms of Energy Loss

Quality factor is always a critical parameter for any type of vibratory MEMS device, and is inversely proportional to the rate at which energy escapes the resonant system. As such, it determines the frequency resolution, stability, and gain with respect to external forcing. While there are many contributing factors to the energy loss of any given resonator, the dominant sources are largely determined by the operational frequency.

MEMS resonators have many applications, ranging from switches and signal processing to clocks and sensors. Because the design of such structures can be radically different, they are typically separated into two terms: traditional MEMS and RF MEMS. Traditional MEMS tend to operate in the kHz and possibly MHz frequency spectrums and can be focused on a variety of applications, including inertial sensors. RF MEMS operate in the high-MHz to GHz frequency spectrums and are focused on maximizing frequency and stability for signal processing or timing applications.

Because maximizing both frequency and Q-factor are important for RF MEMS, the figure of merit of these devices is typically referred to as the $f \cdot Q$ product. For inertial sensors, low frequencies are desirable for maximizing the mechanical scale factor of the device, making $\tau$, or the damping constant, a tempting analogy, which is proportional to $Q/f$. This is not a completely fair analogy, for while it appears correct on paper, the truth is that environmental vibrations can strongly hinder the performance of inertial sensors, and by lowering the operational frequency of the device in order to maximize $\tau$, the designer is hindering the environmental robustness of the device. For this reason,
the frequency of MEMS gyroscopes are typically fixed to roughly 10 kHz; however, this can vary based on application. Nevertheless, maximizing Q-factor without substantially reducing frequency is a desirable trait.

For this work, because these resonators utilize relatively low frequencies, their dominant sources of energy loss are not the same as typical RF MEMS. At higher frequencies, a large number of additional energy loss mechanisms come into play, which include surface conditions, material defects, and phonon-phonon interactions. These mechanisms will not be considered. In this work, the main frequencies of interest occupy the Very and Ultra Low Frequency bands (0.3 kHz to 30 kHz), borrowing radio spectrum terminology. Within these frequency bands, the most critical sources of energy loss have been both theoretically and experimentally verified to be viscous damping, substrate losses, asymmetry losses, thermoelastic damping, and electronic losses [100]. Each one of these loss mechanisms will be discussed below.

3.3.1 Viscous Damping

Viscous damping is energy loss due to physical collisions with air molecules. When thinking of air as a fluid, a vibrating structure must push past molecules to deflect. This can arise primarily due to two mechanisms: squeeze film damping and slide film damping. Squeeze film damping occurs when the vibrating structure is required to push through and deflect air molecules, while slide film occurs when the structure simply slides along a viscous surface. For capacitive MEMS, squeeze film frequency dominates due to the large number of parallel plates electrodes. While this mechanism can be removed by evacuating the cavity which contains the resonator through exposure to vacuum, it is
typically the greatest source of energy loss for structures exposed to air.

The presence of viscous damping is well documented and analyzed. While many analytical expressions exist to model this phenomenon, many are inconsistent between experiments [101]-[108]. For this reason, it is typically beneficial to experimentally measure the damping versus pressure relationship for a given system instead of relying upon existing models.

3.3.2 Substrate Losses

Substrate loss is energy loss due to externally propagating stress waves through the substrate of the resonator. As a one-dimensional resonator vibrates, the vector of momentum oscillates in magnitude along a single line of motion. This oscillating momentum causes deflection in the substrate, producing stress waves that expand outward from the structure. This form of energy loss can be mitigated through the use of anti-phase resonances; however, a certain amount of energy loss can remain due to any unbalanced components. This residual energy loss is typically inconsequential at low frequencies, but increases in magnitude as the oscillating frequency increases due to design modifications.

One method of analyzing a resonant structure for substrate losses is through the use of finite element modeling, with an infinite boundary condition on the substrate. The chosen boundary condition must be capable of perfectly attenuating any acoustic waves, without reflection back into the model. Such an absorption layer is also known as a perfectly matched layer (PML). The perfectly matched layer boundary condition was first introduced by Berenger for the purpose of studying electromagnetic waves [109], and was later adapted for elastodynamic systems [110], [111]. It is now a standard feature of
Figure 3.15: Model of two-mass gyroscope used for finite element modeling of anchor loss, consisting of the device, substrate, and perfectly matched layer (PML) for the purpose of absorbing strain energy.

Figure 3.16: Substrate damping versus number of mesh elements for the in-phase and anti-phase mode shapes of Figure 3.15.
current versions of COMSOL Multiphysics, though careful choice of size and attenuation parameters are necessary for successful implementation. It is recommended to tune the parameters such that the PML is capable of absorbing one wavelength of stress, transmitted at the resonance frequency of the analyzed vibratory mode. An example of such a model for a dual-mass resonator is shown in Figure 3.15, with measured Q-factor due to substrate losses versus mesh density shown in Figure 3.16. For this particular design, the in-phase mode is shown to be limited by substrate losses with a Q-factor of roughly 20 thousand, while the balanced anti-phase motion is able to attain a Q-factor of around 2.5 million; approximately 2 orders of magnitude improvement.

3.3.3 Asymmetry Losses

Asymmetry loss is energy loss due to stiffness or mass imbalance in anti-phase resonators. A form of substrate loss, this loss mechanism is only relevant when utilizing anti-phase resonances. Should the anti-phase resonator have any asymmetry in stiffness or mass across the mirrored plane of symmetry, a momentum imbalance occurs which can lead to an increase in energy loss. Asymmetry is always present in the form of fabrication imperfections, so mitigation of these fabrication imperfections may be required, depending on the tolerances of the fabrication method.

This type of energy loss can be calculated using an identical process as shown previously in Equations 2.2 and 2.3, though with the inclusion of damping and assuming external acceleration is negligible [1]. The resulting equations of motion for such as tuning fork resonator is shown in 3.2.
\begin{align*}
\ddot{x}_{an} + \frac{\omega_{an}}{Q_{an}} \dot{x}_{an} + \omega_{an}^2 x_{an} &= \frac{\Delta k}{2m} x_{an} + \frac{F(t)}{2m} \\
\ddot{x}_{in} + \frac{\omega_{in}}{Q_{in}} \dot{x}_{in} + \omega_{in}^2 x_{in} &= \frac{\Delta k}{2m} x_{an}
\end{align*}

When assuming the structure is actuated with an ideal sinusoidal anti-phase motion with constant amplitude, a solution can be formulated for the in-phase motion of the structure, which is provided in Equation 3.3.

\begin{align*}
x_{an} &= A_{an} \cos(\omega_{an} t) \\
x_{in} &= \frac{A_{an} \Delta k}{2m} \left( \left( \omega_{IP}^2 - \omega_{AP}^2 \right)^2 + \left( \frac{\omega_{IP} \omega_{AP}}{Q_{IP}} \right)^2 \right)^{-\frac{1}{2}} \cos(\omega_{an} t + \phi)
\end{align*}

When considering the amount of kinetic energy stored within a resonator and the definition of Q-factor, an equation can be formed for the amount of energy leakage to the in-phase, as shown in Equation 3.4.

\begin{align*}
E_{\text{Stored}} &= \frac{1}{2} m (A_{an} \omega_{an})^2 \\
\Delta E_{in} &= E_{\text{Dissipated per Cycle}} = \int_0^{2\pi/\omega_{an}} m \frac{\omega_{in}}{Q_{in}} \dot{x}_{in}^2 dt \\
&= \frac{\pi A_{an}^2 \Delta k^2}{Q_{in}} 4m \left( \omega_{in}^2 - \omega_{an}^2 \right)^2 + \left( \frac{\omega_{in} \omega_{an}}{Q_{in}} \right)^2 \\
Q \Delta k &= 2\pi \frac{E_{\text{Stored}}}{E_{\text{Dissipated per Cycle}}}
\end{align*}

The rate of energy dissipation for in-phase vibratory modes of resonance is significantly higher than anti-phase resonances due to the lack momentum balance, potentially by several orders of magnitude, as discussed in the previous section. When it is assumed
that the energy transfer between these modal resonances only travels in one direction, which damps the anti-phase resonance, the degree of this energy loss can be calculated as a Q-factor limit of the anti-phase vibratory mode, as determined by Equation 3.5.

\[
Q_{\Delta k} = Q_{IP} \frac{4m^2 \omega_{AP}}{\Delta k^2 \omega_{IP}} \left( \left( \frac{\omega_{IP}^2 - \omega_{AP}^2}{2} \right)^2 + \left( \frac{\omega_{IP} \omega_{AP}}{Q_{IP}} \right)^2 \right)
\]  

(3.5)

Equation 3.5 can be maximized using a number of methods. First, the same principals discussed in Chapter 2 continue to apply: 1) Attempt to minimize the stiffness mismatch, \( \Delta k \), through process optimization, and 2) Maximize the frequency separation between the in-phase and anti-phase vibratory modes of resonance. In addition, the in-phase Q-factor, \( Q_{IP} \), should also be maximized, if possible. Due to the \( \frac{\omega_{AP}}{\Delta k} \) term, a high frequency anti-phase vibratory mode and low in-phase mode are also beneficial for maximizing \( Q_{\Delta k} \); however, such a strategy is not recommended due to the increased influence of external acceleration on the device, as discussed in Chapter 2.

### 3.3.4 Thermoelastic Damping

The nature of thermoelastic damping was first formulated by C. Zener in 1937 [112], [113], [114], and can be described as an intrinsic source of damping due to the coupling between the mechanical stress and thermal gradient of a material. This type of energy loss is not due to internal friction, where heat is generated as particles slide past one another, but is rather only due to stress without breaking any material bonds. A diagram displaying this relationship, through the use of the material properties of the resonator, is shown in Figure 3.17.
When a force is placed on a beam, the imparted stress deflects the structure according to the modulus of elasticity. The change in geometry is then linked to the thermal domain through the coefficient of thermal expansion. This interaction is usually observed in the other direction, where the geometry of a material changes due to a large change in temperature; however, the same principal functions in both directions. A thermal gradient is then formed across the beam as a function of stress, which attempts to equalize itself by conducting across the width of the beam. Any thermal energy that cancels itself in this way is then lost due to the rising entropy, thus damping the mechanical vibrations.

As can be seen, this phenomena is largely dependant upon the material properties of the resonator; however, this source of energy loss can still be minimized through design. First, because there are two energy domains taking part in this, there are two different
critical frequencies: the mechanical resonance frequency, and the inverse of the heat transfer time constant, \( \frac{1}{\tau_h} \). This time constant represents the rate at which the heat dissipates across the width of the beam, the inverse of which has the same units as the angular rate of the mechanical resonance frequency. When these two frequencies are matched, the energy loss is maximized; however, by separating these frequencies, the effect can be minimized. In order to accomplish this, Zener derived a linearized equation for the thermoelastic Q-factor of a simple beam in one of his original works [113]. This equation is reprinted as Equation 3.6.

\[
Q_{TED} = \frac{E\alpha^2 T_0}{C_v} \frac{\omega_n \tau_h}{1 + \omega_n^2 \tau_h^2}
\]

\[
\tau_h = \frac{h^2 C_v}{\pi^2 k}
\]

where \( \alpha \) is the linear coefficient of thermal expansion, \( C_v \) is the specific heat per unit volume, \( k \) is the thermal conductivity, \( T_0 \) is the equilibrium temperature, \( E \) is the Young’s Modulus, and \( h \) is the beam thickness.

Equation 3.6 can be used to compare the performance of different materials in terms of thermoelastic Q-factor. An example of this can be seen in Figure 3.18, which plots the estimated thermoelastic Q-factor versus the mechanical resonance frequency for three common MEMS materials: silicon, quartz, and fused silica. As can be seen, due to its relatively high thermal frequency of approximately 1 MHz for 10 \( \mu \)m beams, the properties of silicon are excellent for low frequency resonators, such as inertial sensors; however, it performs relatively poorly at high frequencies. The thickness of the beam can possibly be used to adjust this thermal resonance, as shown in Figure 3.19; however,
Figure 3.18: Q-factor due to thermoelastic damping, $Q_{TED}$, versus mechanical resonance frequency for three common resonator materials. Beam thickness is 10 $\mu m$.

this also influences the resonance frequency and is very dependant upon the fabrication process.

While Equation 3.6 is excellent for comparing materials or determining trends, it is only quantitatively applicable to simple beam resonators with small amplitudes of motion. For analyzing complex geometries, an improved estimate can be determined through the use of finite element analysis [115], [116]. While this technique is based off of Zener’s equation, and therefore is still only valid for small amplitudes of motion, it not only includes the energy loss of a single beam, but of every stress gradient across the entire device.

An example of such a model is shown in Figure 3.20, which displays the stress and temperature gradients across the beams of a tuning fork resonator. This analysis can
Figure 3.19: Q-factor due to thermoelastic damping, $Q_{TED}$, versus mechanical resonance frequency for three different silicon beam widths.

be performed for each individual mode of resonance. For the device of Figure 3.20, the resultant thermoelastic Q-factor versus the number of mesh elements is provided in Figure 3.21 for both the in-phase and anti-phase vibratory modes. This particular tuning fork has a conventional spring coupling between the proof masses, which does not deflect during in-phase resonance, but does deflect during the anti-phase motion. This additional stress of the anti-phase mode reduces the thermoelastic Q-factor by over 10%.

The variation in the result of Figure 3.21 is not directly linked to the number of mesh elements, but rather the mesh geometry. This is evident by the lack of a directional bias to the result as the number of mesh elements is increased. For each data point, the mesh is randomly generated to consist of a given number of triangular mesh elements, resulting in a variable asymmetry error during this generation process. This asymmetry
Figure 3.20: Stress and thermal distribution of a two-mass gyroscope.

Figure 3.21: Thermoelastic damping versus number of mesh elements for the in-phase and anti-phase mode shapes of Figure 3.20.
is the reason for the small variations in the result; however, the answer consistently hovers around a nominal value, which is the correct solution.

When designing structures to maximize the thermoelastic Q-factor, it is important to realize that there is more to it than simply choosing the right material and mechanical resonance frequency. For complex structures, the thermoelastic Q-factor can quickly drop if stress concentrations are present. As such, it is important to examine the stress distribution across the entire structure for each relevant mode shape. Due to this same principal, the maximum stress in each flexure should also be equated to maximize the thermoelastic Q-factor.

### 3.3.5 Electronic Losses

MEMS resonators are not only mechanical devices, but also electrical. This is one of the greatest advantages of the field, because it allows the structures to be immediately interfaced with electronics for both actuation and detection of the mechanical elements. This factor alone grants many advantages with respect to C-SWaP and measurement sensitivity. However, just as the front-end electronics of vibratory MEMS can force the structure into actuation, it can also damp the structure when this resonance is being detected.

There are a number of methods that can be used to detection the motion of a MEMS resonator, with an introduction to these methods provided in Appendix A. There are two methods which are commonly employed: 1) Transimpedance amplification, and 2) Transcapacitive amplification. A number of tradeoffs exist for the use of either method, one of which is the degree of mechanical damping which is induced. Transcapacitive
amplification can induce a much higher degree of this electrical damping, compared to transimpedance amplification. This is one of the primary reasons transimpedance amplification has been used throughout this work: to maximize the quality factor of the resonator.

When a capacitor is used to detect the motion of a MEMS resonator, the motion of the structure generates a current. This current is then amplified and converted to a voltage so that the signal can be detected and recorded using conventional methods. Depending on how this amplification and converting is performed, it may consume energy from the resonator, effectively damping the mechanical structure. Previous work has identified this effect when a single resistor was used as the sensing element [117], [118], as well as when transcapacitive amplifiers have been used [119]. To date, the electrical damping induced by transimpedance amplification has yet to be identified; however, it is presumed to be low compared to the previous two methods.

Due to the low mechanical damping influence of transimpedance detection, it has
gone largely unnoticed in the literature; however, with the ultra-high Q-factor devices that have been fabricated as a part of this work, this effect can be seen. In order to derive this influence, a schematic of a transimpedance amplifier is provided in Figure A.17. In a similar derivation as the transimpedance amplifier gain of Equation A.26, a new relationship can be formed between the output current of the MEMS resonator, $I$, and the input voltage of the transimpedance amplifier, $V_1$. This derivation is shown in Equation 3.7.

$$\sum i_1 = 0 = I - V_1 C_p s - (V_1 - V_O) \left( \frac{1}{R_f} + C_f s \right)$$

$$I = V_1 \left( C_p s + \frac{1}{R_f} + C_f s \right) - V_O \left( \frac{1}{R_f} + C_f s \right)$$

$$V_1 = \frac{IR_f}{R_f (C_p + C_f (1 + G_{OL})) s + (1 + G_{OL})}$$  \hspace{1cm} (3.7)

The output current of the MEMS resonator can be derived as a linear combination of the structure velocity and velocity of the changing voltage across the capacitor. This relationship is derived in Equation A.18, which is reprinted here as Equation 3.8.

$$I = \frac{dC}{dx} \dot{x} V_C + CV_C$$  \hspace{1cm} (3.8)

where $V$ is the difference in voltage between the proof mass and capacitive electrode. For this term, assume that the voltage of the electrode is always near zero, making the proof mass voltage dominate. This assumption is necessary to avoid a recursive relationship which would otherwise significantly complicate the continued derivation. In addition,
this assumption is valid as long as the electrostatic damping remains low.

The electrostatic force on the MEMS resonator can also be derived as a function of
the voltage across a parallel plate capacitor. This relationship is derived in Equation
A.12, which is reprinted as Equation 3.9 for the sense electrodes. It is important to note
that when using a oscillating carrier voltage, a force is always generated at the sense
electrodes. As long as this frequency is high compared to mechanical resonance of the
structure, it will have little influence on the mechanical oscillation.

\[
F = \frac{1}{2} \frac{dC}{dx} V^2
\]

\[
F = \frac{1}{2} \frac{dC}{dx} (V_C - V_1)^2
\]  \hspace{1cm} (3.9)

\[
F = \frac{1}{2} \frac{dC}{dx} (V_C^2 + V_1^2 - 2V_CV_1)
\]

It is typically assumed that the voltage of the static electrode is grounded by the
transimpedance amplifier, which essentially would set \( V_1 = 0 \). As can be seen from
Equation 3.7, this is not the case. When accounted for these fluctuations of \( V_1 \), Equations
3.7 through 3.9 can be combined to from Equation 3.10. A new term, \( A \), is also defined
for convenience.
\[ A \equiv R_f (C_p + C_f (1 + G_{OL})) s + (1 + G_{OL}) \]

\[ V_1 = \frac{(\frac{dC}{dx} \dot{V}_C + CV_C) R_f}{A} \]

\[ F = \frac{1}{2} \frac{dC}{dx} V_C^2 + \frac{1}{2} \frac{dC}{dx} V_1^2 - \frac{dC}{dx} C V_1 \]

\[ F = \frac{1}{2} \frac{dC}{dx} V_C^2 \]

\[ + \frac{1}{2} \frac{dC}{dx} \frac{R_f^2}{A^2} \left( \left( \frac{dC}{dx} \right)^2 \dot{V}_C^2 + C^2 \dot{V}_C^2 + 2 \frac{dC}{dx} C V_C \dot{x} \right) \]

\[ - \frac{dC}{dx} \frac{V_C R_f}{A} \left( \frac{dC}{dx} V_C \dot{x} + C \dot{V}_C \right) \]

\[ F = \frac{1}{2} \frac{dC}{dx} \left( V_C - \frac{C \dot{V}_C R_f}{A} \right)^2 \]

\[ + \left( \frac{R_f^2}{A^2} \frac{dC}{dx}^2 \right) CV_C \dot{V}_C - \frac{R_f}{A} \left( \frac{dC}{dx} \right)^2 V_C^2 \dot{x} \]

\[ + \left( \frac{1}{2} \frac{R_f^2}{A^2} \left( \frac{dC}{dx} \right)^3 V_C^2 \right) \dot{x}^2 \]

The final forcing terms of Equation 3.10 are separating by their frequency components, with the second term proportional to the velocity of the resonator, and therefore contributing to the energy flow into or out of the mechanical resonance. Please note that the two terms of this component have opposite signs; the first term pumps energy into the resonator, while the second removes it.

When examining Equation 3.10 and considering ideal operational amplifier behavior, the variable \( A \) is assumed to be infinite. Such an assumption would eliminate all but one term of this equation, making it easy to see why this influence has been ignored previously. Regardless, there are still terms that such an assumption would influence more than others. Of the five terms of Equation 3.10, two are divided by \( A \) and two
Figure 3.23: Q-factor versus the amplitude of the carrier voltage for a tuning fork resonator with transimpedance readout electronics with fit using Equation 3.12. Depicts curves for three electronics that have four different feedback resistors: 23.7 Ω (blue), 46.4 Ω (green), 90.9 Ω (red), and 196 Ω (teal).

are divided by $A^2$. While $A$ will not be assumed to be infinite in this case, the value will still be significantly high. This assumption forces the terms divided by $A^2$ to be of significantly less importance and will hence be disregarded. Therefore, there is only one remaining term that influences mechanical damping, which can be converted into an expression for Q-factor, as shown in Equation 3.11.

\[
Q = \frac{\omega m}{c} \\
c = \frac{R_f}{A} \left( \frac{dC}{dx} \right)^2 V_C^2 \\
Q = A \frac{\omega m}{\left( \frac{dC}{dx} \right)^2 V_C^2 R_f} \tag{3.11}
\]
In order to confirm this relationship, the quality factor of a high-Q tuning fork resonator was measured for a variety of carrier voltages and feedback resistances. The relationships between the total device Q-factor, carrier voltage magnitude, and feedback resistance of Equation 3.11 were then combined to form a fitting equation to the experimental data, as provided in Equation 3.12.

\[
Q_{\text{Total}} = \left( \frac{1}{Q_{\text{Limit}}} + \frac{V_C^2 R_f}{B} \right)^{-1}
\]

(3.12)

where \(B\) is a constant representing the unaccounted for components of Equation 3.11.

A plot of the data and fitted equations is provided in Figure 3.23, which shows a good agreement between theoretical and experimental results. When fitting Equation 3.12 to the experimental data, separate fits were determined for each feedback resistor. The fitting constant, \(B\), as well as the pre-existing Q-factor limit, \(Q_{\text{Limit}}\), are provided in Table 3.2 for each data set.

For each data set, the pre-existing Q-factor limit, \(Q_{\text{Limit}}\), remains relatively unchanged at a value of approximately 2.67 million. The fitting constant, \(B\), however, reduces considerably as the feedback resistance increases beyond roughly 50 \(\Omega\). It is currently unclear why this is the case, but is most likely due to one or more of the assumptions.

<table>
<thead>
<tr>
<th>Feedback Resistance, (R_f) ((\Omega))</th>
<th>Fitting Constant, (B) ((10^\text{e}))</th>
<th>(Q_{\text{Limit}}), Millions</th>
</tr>
</thead>
<tbody>
<tr>
<td>23.7 (\Omega)</td>
<td>4.2979(10^\text{e})</td>
<td>2.6736</td>
</tr>
<tr>
<td>46.4 (\Omega)</td>
<td>4.3682(10^\text{e})</td>
<td>2.6698</td>
</tr>
<tr>
<td>90.9 (\Omega)</td>
<td>3.3923(10^\text{e})</td>
<td>2.6685</td>
</tr>
<tr>
<td>196 (\Omega)</td>
<td>1.7269(10^\text{e})</td>
<td>2.6764</td>
</tr>
</tbody>
</table>

Table 3.2: Electronic damping fitting parameters for Figure 3.23.
made during the derivation of the equation. Regardless, this fitting constant remains relative consistence with a value of $4.3\times10$ for feedback resistances less than 50 $\Omega$, which supports the provided theoretical derivation under small parameter approximation.

### 3.4 Experimental Confirmation of Parameters

So far, only analytical and finite element modeling has been used to predict the parameters of the final fabricated devices. Fabrication of the quadruple mass gyroscopes were completed using the in-house, silicon-on-insulator process, as described in Appendix B. Upon completion, devices were tested for experimental confirmation of the modal frequencies, Q-factor, and parasitic motion, as shown below.

#### 3.4.1 Frequency Responses of Modal Resonances

In order to identify the modal resonances of the structure, a single proof mass was actuated and detected, similar to the process modeled in Figure 2.7. A diagram of this interfacing scheme is provided in Figure 3.24 (top-right), with the green arrow representing the forcer and the blue representing detection. Using a constant force amplitude, the frequency of the actuation was then swept from $1500$ $kHz$ to $4500$ $kHz$, with the magnitude and phase responses shown in Figure 3.24. While the frequencies of these resonances are about $15-25\%$ lower than the modeling suggested, the modal order is believed to remain consistent, with the corresponding mode shapes displayed for each resonance peak.

In order to confirm that each experimentally observed resonance peak indeed corresponds to the modeled mode shapes, additional detection electrodes were incorporated on each proof mass. A single mass of the structure was continued to be used for actuation,
so that no changes would be made to the motion of the structure with respect to Figure 3.24, only in how that motion is identified.

The orientation of the detection electrodes on each proof mass were modified to create tailored sensitivity to each predicted mode shape. The experimental results of each of these resonance sweeps are provided in Figures 3.25 through 3.28. For each orientation of the detection electrodes, a single resonance peak appears with a much higher magnitude compared to the other resonances, confirming the modal identification provided by the finite element analysis.
Figure 3.25: Experimental frequency sweep of a Phase 1 QMG by forcing a single mass and detecting all four masses for anti-phase motion (left), phase of this sweep (bottom right), and active electrode diagram (top right).

Figure 3.26: Experimental frequency sweep of a Phase 1 QMG by forcing a single mass and detecting all four masses for double anti-phase motion (left), phase of this sweep (bottom right), and active electrode diagram (top right).
Figure 3.27: Experimental frequency sweep of a Phase 1 QMG by forcing a single mass and detecting all four masses for opposing in-phase motion (left), phase of this sweep (bottom right), and active electrode diagram (top right).

Figure 3.28: Experimental frequency sweep of a Phase 1 QMG by forcing a single mass and detecting all four masses for in-phase motion (left), phase of this sweep (bottom right), and active electrode diagram (top right).
Because this experiment was performed under atmospheric conditions, the associated damping for each mode shape was approximately identical, with Q-factors of about 20. A high level of damping is necessary for this type of experiment, especially for anti-phase resonators where there can be a large Q-factor variability between the vibratory modes. Should a single high-Q mode exist, the force required to actuate such a mode would be substantially lower to avoid collisions. This low actuation force would then render the low-Q modes unobservable. For this reason, should such an experiment be desired, it must be performed prior to vacuum sealing.

### 3.4.2 Q-factor of Modal Resonances

Much like the modal resonance frequencies of the previous section, a similar analysis can be performed for identifying the Q-factor of each mode. First, because viscous damping is the primary source of energy loss for all of the modal resonances, the environmental air pressure must be removed. This is accomplished by placing the device and supporting electronics within a vacuum chamber and reducing the pressure. As the number of air molecules surrounding the device is reduced, Q-factor increases according to Equation 3.13.

\[
Q_{\text{Total}} = \left( \frac{1}{Q_{\text{Limit}}} + \frac{1}{Q_{\text{Viscous}}} \right)^{-1}
\]

(3.13)

where \( Q_{\text{Total}} \) is the measured Q-factor of the device, \( Q_{\text{Limit}} \) is the limit when viscous damping is eliminated, and \( Q_{\text{Viscous}} \) is the Q-factor due to viscous damping.
Figure 3.29: Amplitude decay measurement for anti-phase Phase 1 QMG motion after given an initial impulse for three different pressure levels. The rate of amplitude decay is directly related to the damping in the system (τ or Q-factor).

By modifying the polarity of the actuation and detection electrodes on each proof mass to represent the mode shape in question, Q-factor can be measured at multiple pressure levels within the vacuum chamber. When τ is large (> 10 s), it can be beneficial to use the amplitude decay of the modal resonance as a measurement of Q-factor, as opposed to a frequency sweep. For a device with an operational frequency of 2 kHz, this occurs at a Q-factor of over 50,000. This is accomplished by observing frequency spectrum of the output of the device, giving the device an impulse to actuate the mode of interest, and recording the rate of decay over time. This rate of decay can then be used to calculate the decay constant, τ, from which Q-factor can be derived. Three examples of this type of measurement are provided in Figure 3.29 for the anti-phase mode of the Phase 1 QMG.

For low values of τ, a frequency sweep of the modal resonance can be used instead. Examples of this technique are provided in Figure 3.30 (left). When these measured Q-factors are plotted versus the cavity pressure, along with the data from Figure 3.29, the inverse sum relationship of Q-factor can be observed. In addition, Equation 3.13 can be fit to the data to determine the non-viscous Q-factor, $Q_{\text{Limit}}$, of the analyzed modal
This procedure can be repeated for each of the modal resonances, modifying the polarity of the actuation and detection electrodes for each mode shape in question. The combined result is displayed in Figure 3.31 for the Phase 1 QMG.

From Figure 3.31 it can be seen that the anti-phase modal resonance is capable of a Q-factor in excess of 1 Million, assuming that near $\mu-Torr$ cavity pressures can be achieved. This experimental Q-factor is within 10% of the modeled thermoelastic damping of the Phase 1 QMG, giving strong evidence that the energy loss of the anti-phase resonance is indeed dominated by thermoelastic effects. The non-viscous Q-factor limit of the in-phase and opposing in-phase vibratory modes are the lowest of the four resonances, with values of approximately 5,000. This is primarily influenced by substrate losses, due to modal shapes which are unbalanced in momentum. The double anti-phase resonance, however,
Figure 3.31: Q-factor versus pressure for all four primary mode shapes of a Phase 1 QMG.

is ideally balanced in momentum, leading to a slightly higher Q-factor limit of 15,000. The reason this Q-factor is not high higher is due to asymmetry losses. This resonance uses the external levers as flexing element, which have thick widths and multiple etch holes. Due to the size of these elements, a high degree of asymmetry is present, resulting in a high value of $\Delta k$ and low value of $Q_{\Delta k}$. In this way, the anti-phase resonance of the structure is isolated from the parasitic modal resonances in both frequency and Q-factor.

### 3.4.3 Parasitic Motion of the Device

One of the assumptions for the derivation of the equations of motion of Coriolis vibratory gyroscopes, shown in Equation 3.1, is that the out-of-plane, z-axis motion of the structure is negligible. In order to justify this assumption, it is shown in Appendix A that high aspect ratio fabrication is a general design strategy to avoid this additional degree of
freedom. While it is generally recommended to pursue aspect ratios in excess of 10 to eliminate this axis from the equations of motion of the system, this strategy only applies to large amplitudes of motion. Despite such fabrication, it is still possible that the z-axis motion may be comparable to the displaced caused by the Coriolis force during measurement, which may influence performance.

One of the ways this can occur is due to electrostatic levitation [120]. The final fabricated device has a solid substrate beneath it and only empty space above it. When forcing the structure into resonance electrostatically, an electric field is generated between the stationary electrode and the proof mass, which induces a force between the two structures. While this force is dominate in the two dimensional plane of interest, when considering the three dimensional space in which the field is generated, there is a degree of field asymmetry due to the solid substrate only being beneath the device. This asymmetry induces an additional force in the z-axis, which is otherwise known as electrostatic levitation.

One of the benefits of the quadruple mass design is that the proof mass, which transfers the Coriolis force between the two axes of motion, is partially decoupled from each of the shuttles where the electrodes reside. This partial decoupling allows a single axis of motion to transfer between the shuttle and the proof mass with extremely low loss, yet attenuates motion along the other two axes. This attenuation is much greater for the undesired in-plane axis, however, it also applies to the out-of-plane motion.

In order to experimentally verify this effect and determine if this out-of-plane motion has the potential to influence device performance, the amplitudes of motion of the struc-
tecture were measured. In-plane amplitudes were measured with a microscope and image processing, while out-of-plane amplitudes were measured with a laser doppler vibrometer.

To confirm that the vibrometer was only measuring out-of-plane motion, and not an in-plane projection due to reflections, a single mass of the gyroscope was actuated and observed using the vibrometer at two different device orientations: 1) As flat as possible, and 2) With an induced tilt of approximately 14°. The experimental setup is shown in Figure 3.32 (top-left). The experimental results from the vibrometer are shown in Figure 3.32 (bottom) for both orientations, along with a 95% confidence interval. The results are nearly identical, with a reduction in amplitude for the tilted data equal to the projection of the out-out-plane motion onto the sensing vector of the vibrometer. Apart from this,
the only consequence from the large tilt in the device was a larger degree of error, most likely due to a decrease in the amount of reflected light for measurement.

With the experimental procedure confirmed, the in-plane and out-of-plane displacements of the structure were measured at various locations across the device. With a maximum in-plane motion of \(0.5 \, \mu m\), an out-of-plane motion of \(5 \times 10^{-11} \, m\) was measured at the shuttle of the forcing electrodes. This can be compared to an amplitude of \(2 \times 10^{-12}\) at the cross-axis shuttles, and a noise level of \(5 \times 10^{-13}\). This data is provided in Figure 3.33.

This data shows that the out-of-plane displacement of the shuttles which contain the forcing electrodes is approximately four orders of magnitude less than the in-plane mo-
tion. Assuming this is due to electrostatic levitation, vacuum sealing should even reduce this further because less driving force will be required for a similar in-plane amplitude. As this motion attenuates between the driving shuttles and the proof mass, then again between the proof mass and the sensing shuttles, there is nearly another two orders of magnitude reduction in displacement, which is only barely above the resolution of the measurement technique. This data shows that shuttles are not only an excellent method of decoupling in-plane motion, but also out-of-plane. As such, whenever possible, use separate shuttling structures when forcing and detecting the motion of a proof mass along a single axis of motion.

3.5 Instrumentation and Performance

Vibratory inertial sensors consist of more than just a mechanical element. While the design of the mechanical resonator does typically limit performance because it serves as the initial transformer between the inertial input and mechanical displacement, signal condition and control loops are also necessary components that must be optimized to prevent the deterioration of performance. In addition, compensation algorithms can also be incorporated into the signal output to reduce a number of deterministic drifts, such as those due to temperature or acceleration. These additional factors will be discussed in this section.

3.5.1 Front-End Electronics

As shown in Appendix A, electrostatic transduction is a useful method to convert between the mechanical and electrical domains of a MEMS resonator; however, a challenge
exists when attempting to establish a control loop for maintaining a constant amplitude of resonance. Electrostatic actuation converts an electrical voltage to mechanical force, while electrostatic detection converts a mechanical velocity to electrical current. In addition, conventional electronics operate by detecting or creating a variable voltage, not current. For this reason, at the very least, additional electronics are necessary for converting the output current of the device to a voltage, either for feeding back to the driving electrodes for self-resonance actuation, or digitizing. In addition, depending on the actuation scheme of the device, injecting a DC voltage to the AC signal may also be necessary before being sent to drive the structure. For these reasons, front-end electronics are required.

A photograph of the front-end electronics used for this testing is shown in Figure 3.34. This figure depicts a MEMS gyroscope in a dual in-line (DIP) package, inserted into a Printed Circuit Board (PCB) which contains the actuation and detection signal conditioning electronics. A number of Bayonet Neill-Concelman (BNC) cables are then attached to the PCB and fed out to a lock-in amplifier from Zurich Instruments for digital control of the device. The PCB has been placed within a metal housing, which is currently open to view the inside. This housing behaves as a Faraday cage, reducing the influence of external electric fields on the output of the gyroscope, before the information can be amplified and digitized. Because the output current of the device is extremely low for small rotations, great care must be taken to preserve the integrity of the signal before it can be processed for increased robustness.

For additional information on the details of the front-end electronics, please see Ap-
3.5.2 Control Loops

The control of the gyroscope is completed through the use of a lock-in amplifier from Zurich Instruments. This amplifier provides the necessary functionality for implementing an open-loop control scheme for the gyroscope, as depicted in Figure 3.35. In this control scheme, the driving axis of the gyroscope is actuated into a constant amplitude of motion, through the use of two feed-back loops: 1) Amplitude control, and 2) Phase control. The two control loops are necessary for decoupling and independently controlling the two degrees of freedom of the actuation signal.

Each control loop consists of at least three stages: 1) Demodulation, 2) Proportional-Integral-Derivative (PID) Control, and 3) Modulation. The demodulation step first splits the output of the drive-mode into two, linearly independent components. This is accomplished by multiplying the output of the drive-mode by the frequency of oscillation, then
Figure 3.35: Schematic of the open-loop gyroscope control scheme.

...low pass filtering the result. Depending on which parameter that is being isolated, the multiplied frequency may either be in-phase with the gyroscope (for amplitude detection), or 90° out-of-phase (for phase detection). After demodulation, the phase should be zero; however, to make sure it remains as such, it is fed into a PID controller, the output of which is the control signal necessary to maintain the input signal at zero.

A similar process occurs for the amplitude measurement; however, because a constant, non-zero amplitude is required, this signal must be subtracted from the value of the set point before being fed into the PID controller. The phase control signal is then fed into a Voltage Controlled Oscillator (VCO), which creates a high stability frequency reference which is used for both driving the structure, as well as the demodulation steps. The amplitude control signal can then be modulated by the frequency reference and used to drive the structure.

Because this is an open-loop control scheme, the sense mode is left free to oscillate...
due to the forcing it experiences, preferably from the Coriolis effect. This motion is then
demodulated, resulting in the rate output of the device.

3.5.3 Bias and Scale Factor

Once the control loops have been established, the device is now capable of operation.
The first parameter that must be identified is scale factor, or the conversation between
the voltage output of the device and the applied inertial rate of rotation. While it is
possible to analytically calculate this conversion, a number of variabilities exist over the
course of fabrication, making experimental characterization still a necessary process. To
complete this step, the gyroscope must be mounted on a rate table, capable of producing
a rate stability in excess of that of the device. Multiple sinusoidal rates can then be
applied and the amplitude response of the gyroscope recorded. A line can then be fit
to this data, with scale factor represented as the slope, and the zero offset of the device
represented as the linear offset of this fit.

Once scale factor has been identified, the bias of the gyroscope can be measured.
This involves leaving the gyroscope under static conditions and recording the output
over time. Once a significant data set has been recorded, the Allan variance of this data
can be calculated according to Equation A.28. A conceptual output of this calculation
is provided in Figure A.31, along with additional information on this type of analysis.
Concerning typical gyroscope performance, there are at least three parameters that can
be derived from this plot: 1) Angle random walk, or white noise, 2) Bias instability, and
3) Rate random walk, or red noise. By fitting segments of the Allan variance plot to
slopes of $-\frac{1}{2}$ and $+\frac{1}{2}$, the angle random walk and rate random walk can be identified,
Figure 3.36: Example of an Allan variance curve of the zero rate output of the gyroscope. The lowest value on the curve is the bias instability.

An example of an actual Allan variance plot of a Phase 1 QMG is provided in Figure 3.36. The solid line represents the measured value, while the dotted lines represent the region of error. Because there are less data points available for larger cluster times for a fixed data set, the error expands as cluster time increases. The peak in this plot at a cluster time of approximately 0.5 s is most likely due to low-pass filtering of the data during control implementation, and conceptually should continue upwards as part of the $-\frac{1}{2}$ slope of white noise.

For this particular gyroscope, the scale factor was previously calculated to be approximately $1 \,(\circ/s)/\mu V$, resulting in a bias instability of $90 \, \circ/hr$. During device fabrication, there is a high degree of variability between devices, primarily due to two reasons: 1) Academic fabrication equipment, making process recipes drift quickly over time, and 2)
A large device area of roughly $8 \times 8 \text{ mm}$. This variability results in the fabrication of devices with a wide range of performance metrics. While this particular device is of marginal performance, higher performance devices have also been fabricated as a part of the same fabrication batch, with innate bias instabilities of $0.2 \, ^\circ/hr$ [8].

### 3.5.4 Frequency Stability and Phase Drift Compensation

Frequency and phase fluctuations ($\beta(t)$ and $\phi(t)$) are indistinguishable from one another when attempting to measure an instantaneous frequency or phase ($\beta(t) = \dot{\phi}(t)$). In this way, when attempted to demodulate the amplitude output of an oscillator by its frequency, variations within this frequency can directly translate to a phase error. Considering the fast sampling times of most inertial sensors, this frequency error is typically dominated by white noise. When frequency white noise is integrated to represent a variable phase delay, it develops a red noise characteristic and has the potential to drift far from its initial value due to random walk. This relationship is described by Equation 3.14.

As this phase delay of the rate signal demodulation shifts, the quadrature signal leaks into rate signal output, creating a drift in the zero rate output of the sensor. Assuming small phase deviations, the induced drift in the zero rate output is linear with respect to the phase drift, the rate of which is a function of the magnitude of the quadrature signal.

Due to the red noise characteristic of this effect, it appears in the Allan deviation plot of the rate signal as a noise source with a positive slope of $+\frac{1}{2}$.

\[
\phi(t) = \int \beta(t)dt \\
\phi(t) = \int \beta_A t^{-\frac{1}{2}} dt = 2\beta_A t^{\frac{1}{2}}
\]  

(3.14)
Figure 3.37: Allan variance curve of the zero rate output of the gyroscope. $-\frac{1}{2}$ and $+\frac{1}{2}$ fits are shown as dashed red lines.

where $\phi(t)$ and $\beta(t)$ are the time-dependant phase and frequency variations of the resonator, respectively, and $\beta_A$ is the magnitude of frequency white noise.

In order to test this hypothesis, a single Phase 1 QMG was used within the open-loop control scheme, as shown in Figure 3.35. Multiple parameters were then recorded for the device over a 24 hr length of time, which is long enough to begin to observe red noise in the rate output of this particular device, with a high degree of accuracy. The recorded parameters included the rate output of the device, the quadrature output, as well as the oscillating frequency, as defined by the phase-locked loop. Allan deviation analysis was then applied to the rate output of the device, as well as the oscillating frequency, with examples provided in Figures 3.37 and 3.38. By fitting $-\frac{1}{2}$ and $+\frac{1}{2}$ slopes to the
Figure 3.38: Allan variance of the gyroscope frequency, as determined by the place-locked control loop. $-\frac{1}{2}$ fit is shown as a dashed red line.

To achieve additional data points, the parameters of the PID controller in the phase-locked loop were manipulated to increase the frequency noise of the oscillator, without changing any other parameters. The previously described data collection and analysis was then conducted for each newly acquired data set. By plotting the experimentally determined rate red noise versus the frequency white noise, a clear linear relationship is observed, as shown in Figure 3.39. While only three data points have been provided so far, the resulting linear fit is highly accurate, with an $R^2$ value of over 0.999, demonstrating the dependance of rate red noise on the frequency stability of the oscillator, as predicted. This supplies strong evidence that the physical phenomena behind rate random walk, at
Figure 3.39: Rate red noise versus frequency white noise.

least for this device, is dominated by phase fluctuations due to white frequency noise.

3.5.5 Acceleration Sensitivity and Compensation

Phase fluctuations in the rate signal demodulation is just one of many sources of long-term drift in the output of a sensor. While some of these sources are stochastic, such as the general influence of device aging, many are deterministic, such as the influence of temperature or acceleration. Assuming these deterministic sources are observable, it is possible to algorithmically compensation their influence from the rate output of the device in real-time. One example of this type of compensation is through the use of temperature self-sensing, where the drifting frequency of the resonator is recorded with respect to a high stability, temperature stable clock signal, and used as a measurement of the device temperature. By compensating the rate output of the device with the device
temperature, thermal drifts can be eliminated [142].

A similar mechanism exists by compensating the rate output of the device by the quadrature signal. When demodulating the output of the sense-axis of the gyroscope by the resonance frequency of the device, two output signals are formed, separated from one another by a 90° phase difference: one signal is sensitive to velocity-dependent terms, such as rotation rate, while the second is sensitive to displacement-dependent terms, such as stiffness asymmetry. This displacement-dependent signal is known as quadrature. This method of compensation cancels a number of deterministic influences, including: 1) Phase fluctuations in the rate signal demodulation, 2) Thermal drift of the front-end electronics, and 3) External acceleration.

Phase fluctuations during demodulation are compensated because such drifts result in opposite influences within the rate and quadrature outputs. The demodulation of these outputs occur instantaneously with a fixed phase separation of 90°; therefore, both quantities are influenced by this effect. This method of compensation will work, as long as the magnitude of the quadrature signal remains constant.

Thermal drift of the front-end electronics is compensated due to the fact that both signals are a part of the same output of the device, and are not separated until demodulation occurs. Because both components pass through the same amplification circuits, thermal drifts within these amplifiers will influence both components in a similar way, thus allowing compensation.

Finally, external acceleration is compensated in much the same way as the thermal drift of the electronics. External acceleration can induce capacitive gap changes within
the detection electrodes on the chip. While a high in-phase frequency and differential electrodes can help reduce this effect, there is always a small displacement induced, and differential electrodes do not compensate for nonlinear effects due to changing gap size. However, this influence appears in both the rate and quadrature signals equally, allowing the compensation of this effect.

The use of quadrature compensation for the elimination of long-term drift in the bias of a gyroscope using Allan variance analysis has previously been demonstrated [8], allowing the flicker noise floor to be identified for an extended period of time. To test the same principal with respect to acceleration, a gyroscope was placed on a dual-axis rate table and bias was measured at multiple orientations with respect to the gravity field,
from −1 to +1 \( g \). For a higher resolution, each data point was averaged for a period of time until the bias instability of the device was reached. The results are displayed in Figure 3.40.

The raw sensor output of Figure 3.40 shows a sensitivity to external acceleration equal to approximately 50 \((°/hr)/g\). Due to the long time period of the experiment, the changing temperature of the device due to the environment created a drift in the sensor output over time. To eliminate this effect, the data was first compensated using temperature self-sensing. Once the influence of the drifting temperature was removed, the quadrature signal was then used to compensate the output, after which, no discernable influence to external acceleration could be identified.

### 3.6 Conclusion

A quadruple mass gyroscope (QMG) has been designed for high performance rate characterization. Over the course of three device iterations, the mechanical structure was optimized while considering a number of parameters for enhanced performance. A few of the primary factors include: 1) Low operation frequency of 2 \( kHz \) for enhanced mechanical scale factor, 2) High in-phase frequency of over 5 \( kHz \) for improved robustness to external acceleration, 3) High frequency torsional modes to reduce cross-axis coupling, 4) Independent actuation and sensing shuttles for each axis of motion to reduce the influence of electrostatic levitation, 5) Differential electrode designs for maximum capacitance and symmetry, and 6) Gap adjustments for the reduction of spring undercutting during etching, thus reducing stiffness asymmetries.

Maximizing the quality factor of the modes of operation was also a priority for improv-
ing both the mechanical scale factor, as well as the reducing the white thermomechanical noise of the system. By identifying and modeling the primary energy loss mechanisms for devices within the frequency spectrum of interest, these mechanisms could be minimized by guiding choices during the design and fabrication of the sensors. For the given structure, it was determined that the critical energy loss mechanisms included: 1) Viscous damping, 2) Substrate losses, 3) Asymmetry losses, 4) Thermoelastic damping, and 5) Electronic losses. This resulted in quadruple mass gyroscopes with operational quality factors in excess of 1 Million, which were limited by the fundamental thermoelastic damping of silicon. In addition, the quality factor of each of the parasitic modal resonances were suppressed using a combination of substrate and asymmetries losses, by design.

Front-end electronics and open-loop control was used to characterize the device responses. A compensation algorithm was also developed by using the quadrature response of the gyroscope. Just as frequency self-sensing can eliminate temperature drifts within the rate output of vibratory gyroscopes, quadrature compensation has been shown to be capable of eliminating several different influences of long-term drift. These factors include: 1) Phase fluctuations during the demodulation of the rate output, thus removing rate random walk, 2) Thermal drift of the front-end electronics, and 3) External acceleration.
Chapter 4

High-Q Frequency Modulated Accelerometer

4.1 Introduction

While conventional MEMS accelerometers have continued to increase in performance over the years, a fundamental limit is quickly approaching, which may prevent their use for inertial navigation applications. For navigation, not only is a high resolution required for precise measurement, but also a wide range for robust performance; the ratio of these two metrics is known as the dynamic range of the sensor, and can be expressed in either $\text{ppm}$ or $\text{dB}$. Because conventional MEMS accelerometers convey this information through the use of a voltage output, the voltage source must have a stability less than or equal to that of the measurement. Aggressive voltage source stabilities are on the order of 10 $\text{ppm}$, indicating a corresponding limit to the dynamic range of the sensor. Other challenges also exist for these devices, such as a fundamental trade-off between the mechanical scale factor and bandwidth, as well as competing packaging requirements, as compared to high performance gyroscopes.

By converting to a frequency modulated detection scheme, each of these challenges
can be solved or mitigated. As compared to the aggressive voltage source stabilities of 10 ppm, frequency references exist on the order of 10 ppb, expanding this range by three orders of magnitude. These types of devices are also enhanced by vacuum packaging, allowing the co-fabrication of high-performance gyroscopes for stable sensor alignment. Despite these benefits, there are still a number of challenges before the accepted use of these types of devices.

One of the main challenges is frequency drift with respect to temperature. The stiffness of silicon has a high temperature sensitivity, forcing frequency shifts to be the result of not only acceleration, but also temperature. To overcome this challenge, two accelerometers of opposite polarity are fabricated on each die. By subtracting their measurements, common-mode frequency shifts, such as those due to temperature, can be canceled, resulting in a pure measurement of acceleration. Conversely, their measurements can be summed to cancel the frequency shifts due to acceleration for a die-level measurement of temperature, as well.

In this chapter, the development of a dual, tuning-fork FM accelerometer is discussed. The design is capable of canceling the thermal sensitivity within the frequency measurement, while also attaining a high quality factor for enhanced frequency resolution. A tunable scale factor is demonstrated, along with design modifications for maximizing the mechanical scale factor. In addition, a specially designed device was able to achieve what is believed to be the highest demonstrated Q-factor and decay constant for a MEMS device, with values of 2.3 million and 1300 s, respectively.
4.2 Motivation

Conventional MEMS accelerometers have proven themselves for a variety of commercial applications, including both consumer and automotive. As shown in Figure 1.5, state of the art inertial measurement units (IMUs) which utilize MEMS technology for acceleration detection have comfortably reached bias repeatabilities of 10\ mg, along with scale factor stabilities of 200\ mg. According to Table 1.5, these metrics are adequate for use in even automotive stabilities systems; however, performance must continue to improve before the use of these devices in inertial and navigational applications. Table 1.5 shows that a resolution of at least 4\ \mu g is required for these purposes, therefore bias must be improved by over three orders of magnitude.

For amplitude modulated devices, this is a challenging task, particularly when you also include the desired range of 1\ g, which corresponds to a required dynamic range of 4\ ppm. One of the many challenges faced for meeting his goal is simply due to the stability of available voltage sources. To produce an analog voltage output which corresponds to a measurement stability of 4\ ppm, the voltage which carries this information must also have at least a stability of 4\ ppm, which is a substantial challenge [122]. Nearly all commercially available, high stability DC power sources are unable to achieve this goal. Most popular brands are limited to stabilities of over 100\ ppm [123], with specialty manufacturers able to approach 10\ ppm [124]. While specialty circuits have been designed to produce voltage references as low as 1\ ppm, [125], [126], the DC bias is locked and can not be changed. Without advances in the stability of voltage sources, amplitude modulated accelerometers may be able to barely meet the navigational grade requirements of Table 1.5, but a firm
Table 4.1: Amplitude modulated versus frequency modulated detection schemes.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Amplitude Modulation</th>
<th>Frequency Modulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Typical Stability</td>
<td>10000 ppb</td>
<td>10 ppb</td>
</tr>
<tr>
<td>Dynamic Range Limit</td>
<td>100 dB</td>
<td>160 dB</td>
</tr>
<tr>
<td>Packaging</td>
<td>Low-Q</td>
<td>High-Q</td>
</tr>
</tbody>
</table>

boundary is being approached.

One method in order to circumvent this challenge is through the use of a frequency modulated detection scheme. MEMS technology can still be used to fabricate the resonant device, thus retaining the desirable SWaP metrics, but simply the mode of detection is no longer based on the amplitude of an output voltage, but rather a quasi-digital change in frequency. Furthermore, frequency references are currently available with stabilities on the order of 10 \( \text{ppb} \), compared to the typical voltage stabilities of 10 \( \text{ppm} \). In respect to the maximum potential dynamic range of the sensor, this corresponds to 100 \( \text{dB} \) for amplitude modulated devices versus 160 \( \text{dB} \) for frequency, as shown in Table 4.1. While this technique clearly offers the potential of a major improvement in device performance, there are still additional challenges that remain for these FM systems.

### 4.2.1 Amplitude versus Frequency Modulation

When transmitting information using a carrier frequency, there are two general methods: amplitude modulation (AM) and frequency modulation (FM). For an amplitude modulated signal, the signal and the carrier frequency are multiplied, resulting in a pulsed waveform as shown in Figure 4.1 (right). The signal information is carried on the amplitude of the output voltage and can be extracted using signal processing, as discussed in the electromechanical amplitude modulation section of Appendix A. While not all tra-
Additional accelerometers use this technique to transmit information, they do rely on the direct detection of proof mass motion for the measurement of the applied acceleration, forcing a voltage output signal. A basic schematic of this type of accelerometer is shown in Figure 4.1 (left), where mass $m$ is suspended by flexure $k$ with deflection state variable $x$.

A frequency modulated accelerometer is very similar in design to an amplitude modulated device, only with the addition of an electrostatic spring, as shown in Figure 4.2 (left), where the potential across the electrostatic spring has a fixed value of $V$. The accelerometer is then driven with a small, fixed amplitude of motion, and is allowed to
drift from its nominal value as a result of applied acceleration. This causes the gap of the electrostatic spring to change with respect to the external acceleration, thus imparting a variable stiffness to the device, as determined by Equation A.17. This stiffness changes the resonance frequency of the device, which then carries the measurement information, as shown in Figure 4.2 (right). While this method does not fully escape the voltage stability limitation of amplitude modulated devices due to the applied voltage of the electrostatic spring, other methods exist which can cause acceleration sensitive stiffness changes, such as axial loading of the flexures [127].

4.2.2 Comparative Performance Analysis

Beyond the voltage stability limitation, amplitude modulated accelerometers have a number of additional challenges. First, because it is a single degree of freedom oscillator where the amplitude conveys the measurement, a tradeoff exists between the scale factor and bandwidth of the device. As shown in Appendix A, device operation assumes relatively static input accelerations to avoid the mechanical resonance of the device. Assuming this is true, the mechanical scale factor is simply \( \frac{m}{k} \), as shown in Equation 4.1.

\[
x = \frac{m}{k}a = SFa
\]  

(4.1)

When measuring a dynamic acceleration, this assumption holds true for input acceleration with frequencies below roughly 50% (47.6%) of the natural resonance for any mechanical resonator, as shown in Figure A.9. This bandwidth can be expanded to just barely higher than the resonance frequency of the device by optimally damping the
oscillator; however, this method is still limited by the low pass filtering effects of the mechanical system. By changing the stiffness of the device, a designer must choose between a higher mechanical scale factor to achieve higher resolution, or a wider bandwidth for robust performance.

By damping the mechanical system, not only can a designer potentially double the bandwidth of the device, but also guard against false measurements. Should an accelerometer with a high Q-factor receive an impulse due to shock, the resonator will begin to oscillate at its natural resonance, supplying a false measurement for a length of time on order with its mechanical decay constant. For these reasons, damping amplitude modulated accelerometers is ideal; however, a challenge arises when attempting to create IMUs. This packaging requirement is opposite to that of high performance gyroscopes, which ideally require high vacuum and Q-factor for enhanced sensitivity, making single-chip IMU integration difficult, if not impossible. Single-chip integration is a desirable quality for IMUs for retaining sensor alignment over time and environment. Currently, IMUs assemble multiple packaged sensors as rigidly as possible to preserve alignment; however, a major source of error remains due to thermal expansion mismatches of the packaging materials, or simply deflection due to external shock. Such changes in alignment can drastically degrade IMU performance, which becomes even more challenging as sensor resolution continues to decrease.

Frequency modulated accelerometers offer solutions to each of these challenges. First, while the bandwidth of the FM devices is still limited by the resonance frequency of the device, it is now decoupled from the mechanical scale factor. The voltage of the electro-
static spring now determines the scale factor, which also has an added benefit of allowing for in-run scale factor tuning. In addition, vacuum packaging of the accelerometer now enhances performance by improving frequency resolution. Not only does this allow for single-chip IMU integration, but it is an additional avenue for increasing performance by maximizing the quality factor of the device. Frequency modulated sensor architectures are also known to provide additional robustness against mechanical and electromagnetic interference [128], [129].

Despite these benefits, there are also a few new challenges that arise due to FM operation. The first is Q-factor. While enhancing the quality factor can improve performance, the low quality factors of typical MEMS fabrication can quickly remove any competitive edge. Another major challenge is temperature drifts. The Young’s modulus of silicon has a strong temperature dependency, making these types of accelerometers also thermometers. Previous attempts at FM accelerometer design have had their performance limited by these dependencies [127], [130], making these factors the main targets of this work.

4.3 Structural Design and Dynamics

In order to meet these challenges, a tuning fork FM accelerometer was designed. It is known from Chapter 3 that substrate damping can be a substantial loss mechanism for MEMS resonators, and that this can be minimized by creating vibratory mode shapes which are balanced in momentum. In this way, a high Q-factor anti-phase mode can be formed for the purpose of creating a high resolution frequency output of the device, while retaining a relatively low Q-factor in-phase mode that is sensitive to changes in acceleration. For this design, acceleration induces an in-phase displacement, which changes the
gaps of the tuning electrodes on each proof mass. This gap change induces a stiffness change, as shown in Equation A.17, which influences both the in-phase and anti-phase resonance frequencies. In this way, the anti-phase resonance frequency is sensitive to acceleration, yet retains a high quality factor for enhanced resolution of the frequency output.

This design, however, only solves one of the proposed challenges, leaving temperature drifts to still obfuscate the output of the device through the innate thermal coefficient of frequency of silicon. In order to decouple the frequency shifts in the device due to temperatures from that of acceleration, two FM accelerometers are used, each with tuning plates oriented in opposite directions. In this way, positive acceleration causes the tuning electrode gaps of one resonator to decreases, thus decreasing the resonance frequency, while the other to increase, thus increasing the resonance frequency. By subtracting the frequency of one accelerometer from the other, a differential frequency measurement is obtained which is robust to common frequency shifts in both devices, which includes drifts induced from temperature. Likewise, by adding the frequencies of the two accelerometers, differential frequencies shifts due to acceleration are canceled, allowing temperature of the device to be measured.

Two different dual FM accelerometers were designed and fabricated in-house for experimental testing. Images of the structures are provided in Figures 4.3 through 4.6, which include both a complete view of the device (left), as well as an enhanced view of a single mass (right). An additional design was also fabricated using the SOI MUMPS process with a slightly altered design, and will be discussed later in this chapter.
Figure 4.3: Complete FM accelerometer

Figure 4.4: Single mass of a high-frequency with high-frequency and linear scale factor. FM accelerometer with linear scale factor.

Figure 4.5: Complete FM accelerometer

Figure 4.6: Single mass of a low-frequency with low-frequency and linear scale factor. FM accelerometer with linear scale factor.
Figure 4.7: Electrode placement on single mass of an FM accelerometer. The electrode types are mirrored on the second mass.

4.3.1 Capacitive Electrodes

The capacitive electrodes of the device were chosen based on minimizing any unintentional electrostatic tuning effects, other than the isolated tuning plates. It should also be noted that high amplitude sensitivity of the detection electrodes is not a priority considering the frequency output; a high enough signal-to-noise ratio to identify the frequency of the device is all that is required. This was accomplished by using differential comb electrodes for both the actuation and detection, along with single-sided plate electrodes for frequency tuning. All electrode types are placed on each proof mass for balanced actuation and detection.

An expanded view of a single proof mass of this structure is shown in Figure 4.7, which
clearly depicts the locations of the differential comb drives for actuation and detection, along with the tuning plates. This electrode scheme is mirrored for the coupled proof mass, except for the tuning plates, which instead are both oriented in the same direction.

4.3.2 Assumptions

Unlike the QMG of Chapter 3, the multiple vibratory modes of the tuning fork FM accelerometer are critical for operation: the in-phase vibratory mode induces a sensitivity to acceleration, while the high-Q anti-phase vibratory mode provides enhanced frequency resolution. The equations of motion of this system are provided in Equation 4.2.

\begin{align*}
\ddot{x}_{an} + \frac{c_{an}}{m} \dot{x}_{an} + \frac{1}{m} \left( k_{an} - \frac{\varepsilon tw}{(d \pm x_{in})^3} V^2 \right) x_{an} &= \frac{F}{m} \\
\ddot{x}_{in} + \frac{c_{in}}{m} \dot{x}_{in} + \frac{1}{m} \left( k_{in} - \frac{\varepsilon tw}{(d \pm x_{in})^3} V^2 \right) x_{in} &= a
\end{align*} \tag{4.2}

The assumptions of Equation 4.2 are listed below:

1. Linear rigid body dynamics

2. The rigid body does not rotate \((\frac{\Delta k}{k_x} \frac{3}{2} \omega_y^2 \omega_{in}^2 \leq 0.001, \text{ or } \omega_\theta \geq 5 \omega_n \text{ for } \frac{\Delta k}{k_x} \approx 1.5\%}\)

3. Out-of-plane, z-axis motion is negligible (Fabrication aspect ratio of \(\geq 10\))

4. In-plane, y-axis motion is negligible

5. The in-phase and anti-phase motion is only dynamically coupled by the electrostatic frequency tuning

6. Anti-phase displacement is small compared to in-phase displacement during normal operation.

152
4.3.3 Modal Frequency Placement

The choice of the in-phase and anti-phase natural resonance frequency have a critical impact on the performance of the FM accelerometer. There are three major cases to consider: 1) $\omega_{in} < \omega_{an}$, 2) $\omega_{in} \approx \omega_{an}$, and 3) $\omega_{in} > \omega_{an}$.

When $\omega_{in} < \omega_{an}$, the stiffness of the in-phase vibratory mode is reduced, raising the mechanical transformation between acceleration and in-phase displacement. While it may appear this is beneficial for enhancing sensitivity, it must be considered that it is not this displacement that is being measured, but rather the impact it has on the anti-phase frequency through the tuning electrodes. If the displacement was directly being measured, it would be no different than the traditional amplitude modulated accelerometer design. From Equation 4.2 it can be seen that the frequency of the anti-phase mode is a function of both the anti-phase stiffness, as well as the induced electrostatic stiffness. Because the total anti-phase stiffness is a linear combination of these two influences, raising the mechanical component effectively reduces the relational change in frequency, and hence scale factor of the device. However, there is also a benefit of reducing the scale factor nonlinearity across the range of acceleration magnitudes. In this way, placing the frequency of the in-phase vibratory mode below that of the anti-phase is a method of creating an approximately linear relationship between the anti-phase frequency shift and input acceleration. It should also be noted that reductions in the in-phase frequency have the same range and bandwidth limitations as amplitude modulated devices. Figures 4.3 through 4.6 display accelerometers of this design.

Placing $\omega_{in} \approx \omega_{an}$ is another method which maximizes both the transformation from
Figure 4.8: Design comparison between an FM accelerometer with linear (left) and non-linear (right) scale factor, as dictated by mode ordering.

acceleration to in-phase displacement, as well as in-phase displacement to an anti-phase frequency shift. This type of design is able to enhance the scale factor of the device, while also potentially expanding range and bandwidth; however, at the cost of high scale factor non-linearity. A comparison between these two design approaches are shown in Figure 4.8, which displays the frequency shift versus acceleration of both the in-phase and anti-phase mode shapes of two devices of similar designs.

The last case to consider is when $\omega_{in} > \omega_{an}$. While such a design may be ideal for gyroscopes, where minimization of acceleration sensitivity is desired, this is not the case of FM accelerometers. Raising the frequency of the in-phase vibratory mode only reduces the transformation between acceleration and in-phase displacement, reducing sensitivity.
4.4 Finite Element Modeling

Structural designs were created to reflect the chosen vibratory modes of resonance. Before fabrication, these designs were confirmed through the use of finite element modeling to determine the frequencies and shapes of the primary modes of vibration, along with the critical Q-factors of interest. Because the frequencies of interest are similar to that of the quadruple mass gyroscope, similar design methodologies could be used for predicting the final Q-factor. The results of this analysis are presented in this section.

4.4.1 Modal Resonance Frequencies

There are two primary modes of interest for the tuning fork FM accelerometer: the in-phase and anti-phase vibratory modes. The in-phase mode is sensitive to acceleration and induces frequency shifts through parallel plate nonlinearity, while the anti-phase mode has a high Q-factor for enhanced frequency resolution. In order to preserve scale factor nonlinearity, two initial designs were implemented: one with low frequencies for limited range but high sensitivity, and a second with slightly higher frequencies. One of the concerns for the low frequency device was fabrication yield; previously, the lowest frequency device that had been fabricated in-house had a value of 1.5 kHz, making 400 Hz an aggressive goal. As such, the high frequency design was included in this fabrication run to minimize risk. The modeled resonance frequencies of both of these designs can be found in Figure 4.9, along with a visual description of the mode shapes in Figure 4.10.
Figure 4.9: Finite element modeling of both the low-frequency (left) and high-frequency (right) FM accelerometer designs.

Figure 4.10: Description of how the in-phase and anti-phase resonances of the FM accelerometer interact for acceleration detection.
4.4.2 Thermoelastic Damping

According to Zener’s equation, reprinted as Equation 3.6, thermoelastic damping can be minimized by maximizing the frequency separation between the mechanical and thermal resonance frequencies. For the given in-house fabrication process, the thermal resonance of silicon is approximately 1 $MHz$. This leads to a desire of minimizing the mechanical resonance frequency for the purpose of maximizing Q-factor. QMGs of approximately 2 $kHz$ have been shown to be limited by thermoelastic damping, therefore for further improves in Q-factor, this damping mechanism must be continued to be reduced by further reductions in frequency.

For the low frequency design of the tuning fork FM accelerometer, finite element modeling has shown this value to be approximately 5.7 million.

4.4.3 Anchor Losses

While the primary method of reducing anchor loss is to use an anti-phase resonance, this method is not 100% effective. There always remains some degree of stress wave that propagates from the anchors of the resonator into the substrate. This is likely due to the remaining stress distribution imparted to the device anchors from the deflecting springs, though this has not been experimentally verified.

Finite element modeling can be used to assess this influence, and for the low frequency design, the remaining device qualify factor due to anchor loss was determined to also be 5.7 million. Assuming the remaining sources of anchor loss are negligible (viscous, asymmetry, and electrical), designing a resonator to maximize and equate the two unavoidable sources of energy loss is an effective solution at maximizing the total quality factor of the
Using Equation A.10, the total Q-factor for the anti-phase vibratory mode of the low frequency FM accelerometer has been modeled to be 2.85 million. This value is as a result from equal damping contributions from both thermoelastic effects and anchor losses.

4.5 Experimental Results

The described tuning fork FM accelerometers were fabricated using the traditional in-house silicon-on-insulator process discussed in Appendix B. Concerning vacuum packaging, some were enclosed without getter material with Q-factors on the order of one thousand, while others were sealed with getter material with varying degrees of success. In this section, these devices are experimentally characterized for bias, temperature sensitivity, and scale factor.

4.5.1 Instrumentation

As a single-axis resonator driven into an anti-phase resonance of fixed amplitude, the control scheme of the tuning fork FM accelerometer is very similar to that of the drive-mode of the open-loop gyroscope that had been previously discussed in Figure 3.35. The phase-locked loop of this control scheme can also be used as the frequency output of the device. The difference arises when the dual configuration of the accelerometer is assessed. Figure 4.11 depicts a simplified control scheme, where a PLL and AGC each independently control the two different FM accelerometers. The accelerometers are shown with oppositely oriented frequency tuning plates to reflect the opposite response
to acceleration. These two resonance frequencies can then be separated into common-mode and differential drifts by simply summing or subtracting the resonance frequencies of the PLLs. Due to the opposite orientation of the tuning plates, acceleration along the sensitive axis produces a differential shift in frequency, which is decoupled from the common-mode drifts, such as those due to temperature. In this way, the output of the device can be used as both an accelerometer, as well as a thermometer.

### 4.5.2 Dual FM Accelerometer and Thermal Rejection

A demonstration of this temperature-robust differential frequency measurement can be seen in Figure 4.12. The frequency shift of both individual FM accelerometers were recorded for multiple accelerations from 0 to 1 g, and at two different temperatures: 30 and 75 °C. In this diagram, it can be seen that at a given temperature, the resonators have opposite shifts in frequency due to the opposite polarity of the tuning electrodes.
When temperature changes, there is an absolute shift in frequency, however, the differential measurement remains consistent. In this way, the influence of temperature and acceleration can be decoupled from the output of the device.

While Figure 4.12 demonstrates this influence for two static temperatures, there is one assumption that has yet to be discussed: that both resonators are at exactly the same temperature. When this assumption is violated, thermal influences on frequency can enter the differential measurement, thus degrading performance. One of the benefits of using silicon, however, is that it has excellent thermal conductivity, as shown in Table A.2. By including both devices on a single silicon substrate, not only is there excellent alignment stability between the axes sensitive to acceleration, but the high thermal conductivity leads to a high temperature uniformity across the silicon substrate.
This result was experimentally verified by introducing a dynamic temperature to the differential FM accelerometer. The device was placed flat with zero input acceleration and then heated externally to a temperature of 70 °C. The output of both accelerometers was then recorded as the device cooled back down to room temperature, which is displayed in Figure 4.13. When observing the acceleration output of each individual channel, an acceleration drift of over 500 mg is observed, which is directly correlated to the current device temperature. However, because this is a common-mode shift in frequency, it is nearly completely eliminated after subtracting the two outputs. The error in the differential measurement that remained had a root mean square value of less than 1 mg.

For an improved estimate of the bias instability of the device, Allan deviation anal-
ysis was also performed for a dual FM accelerometer that was kept in a temperature stable environment. The results for both a single channel, as well as the differential measurement are provided in Figure 4.14. Bias instabilities of 6 $\mu g$ were obtained for both measurements; however, small thermal fluctuations quickly began to deteriorate the result of a single channel after 20 s. For the differential measurement, this bias instability was maintained for over an order of magnitude longer.

### 4.5.3 Scale Factor Improvements due to Mode Ordering

So far, all of the experimental results have been gathered from the in-house fabricated devices displayed in Figures 4.3 through 4.6. These devices were designed with a linear
Figure 4.15: Frequency shift of a linear FM accelerometer versus acceleration with different tuning voltages, revealing a linear tunable scale factor for the device with maximum magnitude of $\sim 2,000 \text{ ppm/g}$.

scale factor, as shown in Figure 4.8, by placing the in-phase frequency of the device below that of the anti-phase frequency. Through the nature of the electrostatic tuning, this allows for a tunable scale factor by adjusting the voltage of the parallel plate tuning electrodes on each mass; however, the maximum achievable scale factor is limited by the in-phase mode snapping to the tuning electrodes. For the linear designs presented, this scale factor is approximately $4.4 \text{ Hz/g}$. The nature of this tunable scale factor is experimentally demonstrated in Figure 4.15.

In order to further enhance this scale factor, and thus the measurement sensitivity of the device, an alternative design was tested where the in-phase and anti-phase resonance frequencies were nearly equated by design. This was enabled by the anti-phase coupling
Figure 4.16: Frequency shift of a non-linear FM accelerometer versus acceleration, revealing a maximum scale factor of $\sim 20,000 \text{ ppm/g}$.

mechanism presented in Chapter 2, with the slight modification of thinning the internal levers to better match the in-phase and anti-phase resonance frequencies. For a quick turn-around time, the devices were fabricated using the commercial SOI MUMPS process offered by MEMSCAP, with a photograph of the device provided in the Figure 4.16 insert.

Once fabricated, multiple $2g$ tip-over tests were conducted for this device and the data points were experimentally fit to Equation 4.2, as shown in Figure 4.16. Under zero acceleration, the scale factor of the new design improved to $109 \text{ Hz/g}$, which is over a 20 times improvement compared to the previous design. This improvement, however, is at the cost of a high scale factor nonlinearity, which varies from $33 \text{ Hz/g}$ at $+1 \text{ g}$ to $186 \text{ Hz/g}$ at $-1 \text{ g}$.

While under open loop conditions, this variance can be compensated algorithmically through the use of a look-up table. An alternative approach is to simply close the loop.
Table 4.2: Improvement in dynamic range of FM accelerometer as a function of Q-factor.

around the in-phase resonance of the device. Such a technique can be used to stabilize the in-phase displacement at a nominal value, which is controlled by the differential frequency shift of the device. This technique is commonly referred to as force-to-rebalance control.

### 4.6 Q-factor Improvements for Increasing Dynamic Range

The devices used to generate Figures 4.12 through 4.16 were all limited in quality factor by viscous damping. For Figures 4.12 and 4.13, the FM accelerometers which were used for data collection were vacuum sealed without getter material, resulting in Q-factors of 600. Getter material was incorporated in the sealing process for the device of Figure 4.14, however, the getter material appeared to have not been fully activated, leaving the Q-factor of the device at 350,000. Finally, the devices of Figures 4.15 and 4.16 were not even vacuum sealed and were exposed to ambient conditions when testing was conducted.

One of the main advantages of the tuning form FM accelerometer compared to traditional AM designs, is that the fundamental tradeoff between scale factor and range / bandwidth is eliminated. While the range and bandwidth of the new design is still limited by the in-phase resonance frequency, scale factor can be independently enhanced by both matching the in-phase and anti-phase resonance frequencies, as well as by maximizing Q-factor of the anti-phase vibratory mode. After experimentally testing the range...
of the 600 and 350,000 Q-factor devices, the dynamic range was calculated to be 90 $dB$ and 130 $dB$, respectively. This information is also provided in Table 4.2. Should viscous damping be completely removed in these devices, the Q-factor should reach the previously modeled value of 2.85 million.

With continued vacuum sealing, such a device was eventually obtained. Figure 4.17 displays the results of a ring-down experiment for the accelerometer, revealing a Q-factor of 2.34 million and $\tau$ of 1300 s. It is believed that this is the highest Q-factor and $\tau$ ever obtained within a silicon MEMS device, and was enabled by maximizing both thermoelastic damping and residual anchor loss that remained after implementing an anti-phase mode shape. While the dynamic range of this device has not been tested, by projecting the trend established by the lower Q-factor devices, a dynamic range of 150 $dB$ can be expected from this device.

### 4.7 Conclusion

There are a number of challenges that face conventional MEMS accelerometers as their performance continues to improve. A few of these challenges include: 1) Different packaging requirements compared to high performance gyroscopes, making single-chip integration difficult, 2) Limitations in available voltage references, restricting dynamic range due to signal noise, and 3) A fundamental tradeoff between mechanical scale factor and bandwidth. One method to overcome these challenges is through a paradigm shift in sensor operation: convert from amplitude to frequency modulated detection.

In this work, a new differential tuning fork FM accelerometer is presented and proposed as an alternative to traditional AM accelerometer designs. By shifting to a fre-
Figure 4.17: Amplitude versus time of a low-frequency FM accelerometer after being given an impulse. Decay reveals a $\tau$ of 1300 s and Q-factor of 2.3 million.

Frequency output of the device, each of the challenges listed above can be addressed: 1) High vacuum sealing enhances performance, which is identical to that of high performance gyroscopes, 2) Frequency references are available which enable approximately three order of magnitude improvement is sensor dynamic range, and 3) The mechanical scale factor of the device can be improved independently of bandwidth.

Similar packaging requirements as high performance gyroscopes is beneficial because it allows single-chip integration of both types of sensors. Because the sensors share a common silicon substrate, there is lithographic alignment between the sensitive axes of the devices. Furthermore, this alignment is highly stable over shock and temperature due to consistent material properties across the die. This fact is critical for the development of IMUs for inertial navigation, where the requirements of alignment stability become
increasingly strict as the performance of the individual sensors continues to increase.

Frequency stability is also a concern for traditional AM accelerometers. As shown in Table 1.5, navigation grade accelerometers require dynamic ranges of less than 4 ppm, while the most precise of voltage references can only obtain 1 ppm. From this information, while it is true that AM devices may be able to just barely meet this requirement, a fundamental limit is being reached. High stability frequency references, on the other hand, are commonly reported with stabilities of up to 10 ppb, allowing dynamic range to expand by another two to three orders of magnitude.

Finally, AM accelerometers are limited by their resonance frequency: a low resonance frequency enhances sensitivity, while a high frequency enhances bandwidth. Tuning fork FM accelerometers fundamentally decouple this behavior through the use of two independent vibratory modes: one mode which is sensitive to acceleration and limits bandwidth, and a second high-Q mode with can be used as the frequency output. By manipulating the ratio of these two resonance frequencies and maximizing Q-factor of the anti-phase resonance, the mechanical scale factor of the device can be enhanced independently from the bandwidth of the sensor.

Previous attempts at FM accelerometers have faced a number of challenges, primarily: 1) Low Q-factors, and 2) High temperature sensitivity. In the current approach, these challenges are addressed through a high-Q tuning fork design and differential frequency measurement. While it true that the current approach is not completely free from dynamic range restrictions due to voltage stability, design modifications are possible to induce frequency shifts through other means, such as axial loading.
In the end, experimental data has shown that ideal modal placement can enhance scale factor by over 20 times, while temperature-robust dynamic ranges are shown to be in excess of 130 $dB$. A device has also been fabricated with what is believe to be the highest reported $Q$-factor and $\tau$ of a silicon device, with values of 2.3 million and 1300 s, respectively. Projecting this result to dynamic range reveals a value of 150 $dB$. 
Chapter 5

Scale Factor Calibration using an Integrated Torsional Resonator

5.1 Introduction

The long-term accuracy of inertial sensors is a critical component for their use in inertial navigation. As is shown in Appendix A, due to the linear relationship between the input and measured rotational rates, there are two primary factors that determine this stability: bias and scale factor. Because bias is independent of rate, it can be calibrated for any device, as long as the input rate during calibration is zero. The accuracy of the calibration is a function of both the stability of the input rate (zero in this case), as well as the stability of the device. Typically, the bias stability of the device is much worse than the stability of the zero rate input during calibration, therefore there is no need for improving the environmental conditions during calibration.

Calibrating scale factor is much more challenging. Because scale factor is dependant upon input rate, a rate must be given to the sensor during calibration, the stability of which must be equal to or below that of the desired sensor stability. The scale factor stability necessary for inertial navigation is below 1 \text{ ppm}, as shown in Table 1.4. Even
the most sophisticated of rate tables are unable to achieve such performance [131].

One solution is to abandon rate gyroscopes completely and focus on whole-angle operation for direct angle measurement. This technique has different stability requirements which may have simpler solutions. While this is surely an attractive option, it requires a complete shift in sensor operation with additional challenges that are not present for rate gyroscopes.

A second option is to forgo external calibration and create a high stability rate signal on-chip. One method of accomplishing this is by embedding a high Q-factor torsional resonator beneath the gyroscope to supply the rate signal for calibration. This technique requires a fabrication procedure of significant complexity, but has the potential for achieving the stability requirements for inertial navigation. This technique will be explored in this chapter.

### 5.2 Motivation

The scale factor of rate gyroscopes can drift for a variety of reasons. As derived in Appendix A, the primary contributing factors of the mechanical scale factor of rate gyroscopes include amplitude, frequency, and for open-loop operation, Quality factor. Additional electrical contributions include drifting gain and phase delays. These factors can be influenced by a variety of external stimuli. One of the major sources of external drift is temperature. MEMS gyroscopes are influenced by temperature in a variety of ways, a few of which include: 1) Thermal sensitivity of the Young’s modulus of silicon, thus altering the resonance frequency, 2) Thermal sensitivity of the gain in the electronics, and 3) Variability in packaging stress resulting in gap changes of the on-chip capacitive
electrodes, thus altering the amplitude of the device through feed-back control. The combined scale factor error of these effects have been known to be in excess of 1000 ppm/$^\circ C$ [156].

However, temperature is only one of many sources of external drift. Other influences include acceleration, shock, and a variety of sources related to aging. Acceleration can induce similar electrode gap changes as thermal drifts, shock can cause collisions which may redistribute mass to alter frequency, and aging in general consists of a number of factors that can create small variations over time in nearly all of these parameters. To achieve the 1 ppm scale factor stability required for inertial navigation, all of these sensitivities must be considered.

5.2.1 Calibration Techniques

Currently, in-run calibration is typically performed through the use of external sensors and post-processing. By knowing how bias and scale factor drift in response to an external stimulus, the direct measurement of this stimuli using additional sensors can be used to compensate the gyroscope output. For long-term drifts due to aging, the device can then be removed at regular intervals between runs for full recalibration. This full calibration procedure can also be performed prematurely should the device experience inputs beyond its safe range of operation, such as when exposed to a high shock.

This procedure is not ideal for two major reasons: 1) Environmental variations between the gyroscope and externals sensors, 2) The man-hours required for the intermittent device recalibration, and 3) The potential for large drifts prior to recalibration. First, there is no guarantee that the environment is consistent between the gyroscope and ex-
ternal sensors. For a device with a 1000 ppm/°C scale factor sensitivity to temperature, this requires the temperature of the sensors to be within 1 mC of each other to maintain a 1 ppm scale factor stability. This can be a challenge, especially when the minimization of size and weight is a priority. Similar issues exist for other external stimuli, such as acceleration and alignment errors.

The second concern is simply the time required to recalibration the device to maintain a high degree of accuracy. Not only must bias and scale factor be calibrated, but also the sensitivities of these factors to the chosen external stimuli that are being measured, all adding to the overall cost of operation. And despite theses costs, there is still the potential for the third challenge: large drifts prior to recalibration. When relying on intermittent calibration, the time between recalibration is typically chosen using the drift statistics of the entire device population. There is no guarantee that any given device is not an outlier that has drifted substantially away from the norm.

Through the use of the proposed on-chip calibration, each one of these challenges is eliminated. By physically applying a calibration signal to the device while in operation, the scale factor, and any associated drift, can be measured and calibrated in real-time.

5.2.2 Error Model

The proposed calibration flow utilizes a gyroscope that has been fabricated on top of a high-Q torsional resonator of fixed frequency. This frequency is designed to be at the upper end of the bandwidth of the gyroscope to sinusoidally encompass the full range of the sensor, without attenuation. The transmitted rate of this resonator to the gyroscope is a function of both the resonator’s frequency and amplitude. Therefore, rate stability
is maximized when both frequency and amplitude stability are also maximized. These relationships are reflected in Equation 5.1.

\[
\Omega_{\text{Cal}} = \theta \omega_{\text{Cal}} \sin(\omega_{\text{Cal}} t) \\
\Delta \Omega_{\text{Cal}} = \sqrt{\Delta \theta^2 + \Delta \omega_{\text{Cal}}^2} \quad (5.1)
\]

where \( \Omega_{\text{Cal}} \) is the transmitted calibration signal, \( \theta \) is the maximum angle of the torsional resonator, \( \omega_{\text{Cal}} \) is the frequency of the torsional resonator, and terms prefaced by \( \Delta \) are the corresponding stabilities.

There is much ongoing research for maximizing frequency stability, as is discussed in Appendix A, but one well known factor for doing so is raising the Q-factor of the device. The resonator essentially behaves as a bandpass filter, removing any noise beyond its bandwidth. Raising the Q-factor of the resonator narrows the bandwidth, resulting in increased frequency stability. For this reason, vacuum-packaging of the resonator with the gyroscope is beneficial.

While raising Q-factor is an excellent method of raising short-term frequency stability, the resonator is still vulnerable to long-term drifts associated with temperature or aging. For this reason, a high-stability clock signal must also be included in the final product for long-term frequency monitoring and compensation.

In additional to frequency stability, amplitude stability must also be controlled. This is typically performed through the use of automatic gain control, as discussed in Appendix A; however, there are still some limitations to this method. As mentioned earlier in this chapter, variations in gap width, along with thermal sensitivities of the electronic
components all contribute to amplitude drift. To remove these effects, a method of direct amplitude monitoring is proposed, through the use of the sideband ratio technique [162]. While this technique is not well suited for small amplitudes of motion, detection sensitivity rises exponentially for large amplitudes, such as those of the torsional resonator.

As this rate signal passes through the gyroscope dynamics, the output of the device is equivalent to the input rate, multiplied by the real-time scale factor of the device. Through filtering and demodulation at the calibration signal, the final output and stability are reflected in Equation 5.2.

\[
V_{Out} = SF \cdot (\theta \omega_{Cal})^2
\]

\[
\Delta SF = \sqrt{2 \Delta \theta^2 + 2 \Delta \omega_{Cal}^2 + Drift^2}
\]

Figure 5.1: Schematic of the control architecture to be used for scale factor calibration using the co-fabricated gyroscope and torsional resonator. Final output and scale factor error are shown.
With the amplitude and frequency of the calibration rate known, the scale factor of the device can be extracted and used to eliminate any drift in the gyroscope. The complete process is visually represented in Figure 5.1. This method is able to calibrate scale factor with the precision shown in Equation 5.2, though with the drift component removed.

5.3 Co-fabrication of Gyroscope and Calibration Platform

The fabrication flow for this co-fabricated gyroscope and calibration platform was derived from the standard silicon-on-insulator (SOI) process used for traditional vibratory MEMS, as described in Appendix B. In this section, the complete co-fabrication procedure is described, along with a summary of the physical results.

5.3.1 Process Flow

The process begins with an SOI wafer of properties similar to those described for the traditional process in Table B.1. The only variants to these properties include: 1) The device layer coating must now consist of silicon nitride, 2) The handle wafer must have similar doping as the device layer for conductivity, 3) The thickness of the handle wafer is reduced to 300 µm, and 4) The handle wafer surface is now coated with 1.5 µm of silicon dioxide. In addition, an external handle wafer is also fabricated from a 300 µm thick Borofloat glass wafer. The glass wafer was patterned and etching through using conventional water-jet milling from an external vendor.

The continued fabrication process is visually described using Figure 5.2 through the
use of device cross-sections. Once the wafers have been acquired, the second step is to create the front-side metallization for the signal routing traces. This is completed by depositing a blanket coating of chrome and gold across the silicon nitride, patterning the coating using lithography, then etching the exposed metal using liquid etching solutions. Afterwards, the wafers are cleaned, patterned again using lithography, and the exposed silicon nitride is anisotropically etched, thus exposing the silicon device layer.

Next, another layer of silicon dioxide is deposited on the surface of the device layer, patterned, and etched using the mask for sensor definition. This step creates a persistent
hard mask for later etching of the device layer. Once this is finished, the front-side processing is relatively complete and backside processing can begin.

Step four shows the backside silicon dioxide being patterned, etched, then used as a hard mask for etching the platform structure. Once complete, the SOI wafer is only held together by the thin device layer, so great care must be applied to avoid breaking the wafer. The backside silicon dioxide is then stripped using dry etching and the external handle wafer is anodically bonded to the platform, as shown in step five. This substrate provides the structural support of the device during the sensor etching of step six. Finally, the device is diced and released using vapor-phase hydrofluoric acid.

5.3.2 Fabrication

The goal of the fabrication process of Figure 5.2 is to embed a high performance vibratory gyroscope on a high stability torsional stage. Through co-fabrication of the two structures, a high stability lithographic alignment can be created between the two resonators. Maintaining this alignment stability over time is crucial to the calibration potential of the technique, and is one of the major factors justifying the complex fabrication process, as opposed to using an external stage.

A number of challenges exist for the success of this process. First, because the gyroscope and platform are separate resonators with different actuation requirements, both structures must be conductive, but electrically isolated from one another; this is accomplished using the buried oxide layer of the SOI wafer. While there will be considerable parasitic capacitance between these two structures due to the thin silicon dioxide layer, the actuation frequencies are significantly separated, allowing the responses to be filtered
Figure 5.3: Schematic of the signal routing used for the vibratory gyroscope, along with a wire bonding diagram for pulling the signals off-chip.

A second challenge is electrically interfacing the gyroscope. Normally, wire bonding is completed to electrically connect the stationary device electrodes to the bond pads of the package. In this case, however, the electrodes of the gyroscope will now undergo motion as a result of the torsional platform. While this motion may be small, the wire bonds would still add an unknown variable stiffness to the platform actuation, along with the potential to break due to fatigue. To avoid these effects, single-mask electrical routing is performed on-chip, then carried to stationary bond pads on the platform device layer through electrically isolated springs. A diagram of this electrical routing and wire bonding is shown in Figure 5.3. An isometric view of the final co-fabricated structure is provided in Figure 5.4, which displays how these springs are electrically isolated from one another.
Figure 5.4: Model of the final co-fabricated device. The four primary layers include the metal electrodes for signal routing, the gyroscope device layer, the torsional calibration platform, and the substrate.

the rest of the platform, yet still structurally supported by the device layer.

The described fabrication process to meet these challenges involves five lithographic masks: 1) Routing metal, 2) Gyroscope structural layer, 3) Torsional platform structural layer, 4) Bonding area definition, and 5) External substrate structural layer. Each of these layers are shown in Figures 5.6 through 5.10, along with the overlapping geometry displayed in Figure 5.5. The device layer consists of a full Phase 2 QMG, with additional space around the parameter to accommodate signal routing. Additional silicon blocks are included to structurally support the isolated routing springs of the platform layer, with silicon islands around the perimeter so that force can be transmitted through the thickness of the device at the critical anodic bonding locations during attachment of the external substrate. Large opening in the device layer are created to release the gyroscope from the surrounding structure, allowing the attached torsional platform to rotate the gyroscope when actuated, along with granting access to the platform electrodes.
Figure 5.5: Complete device.

Figure 5.6: Routing metal.

Figure 5.7: Gyroscope layer.

Figure 5.8: Torsional platform.

Figure 5.9: Bonding metal.

Figure 5.10: Substrate.
Due to the many overlapping processing steps with traditional SOI fabrication, additional details on the fabrication procedure are provided in Appendix B. The largest divergences for this process occur when bonding the external substrate to the SOI wafer stack, as well as when completing the second DRIE etching step for defining the gyroscope. In addition to this, due to the many masking layers involved, device cleanliness is of critical importance to maintain yield.

Photographs of the wafers and devices during step five of the process are shown in Figure 5.11. In this step, the external handle wafer is bonded to the SOI wafer with etched platform. This step was completed successfully with standard anodic bonding procedures, with the only exception being the removal of some of the backside gold to be later used for die attachment. However, this consequence is not critical for performance.
Figure 5.12: Optical profilometry of the silicon torsional resonator after etching. Due to the large gaps, depth could be monitored between etching to achieve the proper etch depth of 300 µm.

One of the reasons for the reductions in thickness of the SOI handle and external substrate is due to the etching of the device layer. Including the external handle, the total thickness of the co-fabricated structure is just over 700 µm. One of the limitations of the PlasmaTherm FDRIE used for silicon etching is that it can only accommodate wafer stacks of less than 1 mm. After applying a 200 – 250 µm thick handle wafer for additional structural support during etching, the total thickness of the stack just barely fits within this limit.

The clearance of the etching machine, however, is not the only concern for etching the device layer. The etching time and quality are also unknown due to the large thickness
of the device and gaps in the bonded substrate. These factors significantly reduce the thermal conductivity across the width of the device, which can have a strong influence on the etch profile [132]. In addition, due to the large thickness and high doping of the silicon, IR light is significantly attenuated as it passes through the structure, preventing the use of the IR method of monitoring etch rate, as discussed in Appendix B.

With the current fabrication process, not much could be done to enhance the thermal conductivity across the wafer stack, though the use of sacrificial materials was investigated. However, due to the large openings in the structure, etch rate could be monitored using optical profilometry, as shown in Figure 5.12. Once the device layer had been etched and the wafers diced, photographs of the final structure can be seen in Figure 5.13, which were taken before the final silicon dioxide release.

5.4 Platform Stability

As shown in Equation 5.2, the effectiveness of this scale factor calibration method is a direct result of the rate stability of the torsional stage. This rate stability can also be decoupled into two terms: 1) Amplitude stability, and 2) Frequency stability. In this
section, the design of the torsional resonator is qualified through the use of traditional
SOI fabrication. Amplitude and frequency stability are measured, along with an analysis
of any parasitic out-of-plane motion that is produced by the structure.

5.4.1 SOI Implementation

Due to the intensive fabrication process of producing the co-fabricated gyroscope and
calibration platform, additional torsional platforms were fabricated using the traditional
SOI process, as described in Appendix B. Due to the fast production time of using this
method, this allowed the design of the torsional resonator to be experimentally analyzed
prior to the completion of the co-fabrication process. The mask used for the fabrication
of these SOI platforms was identical to that of the one used in the co-fabrication process,
through with the inclusion of etch holes so the device could be released. Perhaps it
should also be noted that due to the large suspended area of the resonator (approximately
$12 \times 12 \text{ mm}$), as well as the low frequencies of motion, electrically grounding the substrate
was imperative for operation.

5.4.2 Rate Stability

Once the SOI platforms were fabricated, the structures were placed in a vacuum chamber
and the non-viscous Q-factor limit was measured at a pressure of 0.1 \text{mTorr}. At the
torsional resonance frequency of 1.2 $kHz$, Q-factor was measured to be 70,000 using the
amplitude decay technique. An experimental plot of the result is shown in Figure 5.14.
It is believed that this Q-factor is limited by a combination of substrates losses, through
unbalanced angular momentum, and asymmetry losses, which induce a small degree of
linear momentum, as well.
Figure 5.14: Amplitude versus time of an SOI platform under vacuum, showing a Q-factor of 70,000.

Figure 5.15: Photograph of a packaged SOI platform (left), along with a schematic of the control of the structure during stability testing (right).

While Q-factor is not a direct variable when determining the rate stability of the resonator, it is related to the frequency stability of the resonator once the device is placed within a feedback loop, such as a PLL. This can be seen through the Leeson effect, which is shown in Equation A.38. Assuming a resonator with a drifting resonance frequency, this type of feedback loop is nearly required for constant amplitude oscillations.

Using the on-chip capacitive electrodes, the platform was then actuated into self-resonance using the combination of a phase-locked loop (PLL) and amplitude gain control (AGC), similar to the control scheme implemented for open-loop gyroscope operation.
A photograph of the packaged device and simplified control scheme is shown in Figure 5.15. The velocity and frequency of the resonator was then recorded over time, with Allan variances of the data shown in Figure 5.16, revealing a velocity stability of less than 8 ppm and frequency stability of 0.4 ppm. When combining these results according to Equation 5.2, should the platform be used for scale factor calibration, an accuracy of less than 20 ppm could be expected, which is largely determined by amplitude stability.

It should be noted that this result was obtained without the use of the side-band ratio algorithm for the direct detection of the mechanical amplitude. With an improved AGC, it is believed that amplitude stability can continue to improve for an even higher accuracy measurement.

5.4.3 Parasitic Motion

Similar to the QMG, there is only one useful vibratory mode of the torsional resonator. While it is possible to design the resonator to allow the mode of interest to have the lowest
operational frequency, additional higher order mode shapes exist that can influence the motion of the stage, and potentially influence performance. In order to investigate these effects, the torsional resonator was first modeled using finite element analysis in three dimensions. The results of this modeling for a low frequency, 200 \( Hz \) torsional resonator is shown in Figure 5.17, which displays the lowest five vibrational mode shapes.

As anticipated, the lowest frequency mode shape is the torsional mode of interest. In addition to this, the next four parasitic modes shape up to frequencies of over 10 \( kHz \) all include an out-of-plane component. This indicates that further increases in the resonator thickness can further stiffen these parasitic resonances; however, the maximum thickness of the platform is still limited by the clearance of the FDRIE silicon etching. Another challenge is that assuming a constant aspect ratio for the etch, this would also require larger minimum gaps in the platform design, thereby also reducing the sense capacitance

---

**Figure 5.17: Finite element modeling of the torsional platform.**
Figure 5.18: Depiction of the phase variation in the torsional resonator when analyzed for out-of-plane motion using a laser doppler vibrometer during in-plane actuation.

and sensitivity of amplitude detection. This is a particular challenge when considering the results from Figure 5.16, where it was discovered that amplitude stability is the primary limit of the scale factor calibration potential. For these reasons, a platform thickness of 300 \( \mu m \) was maintained.

In addition to modeling, the motion of the SOI platform was also experimentally characterized under atmospheric conditions. Similar to the QMG experiments of Figures 3.32 and 3.33, in-plane amplitudes of the platform were measured with a microscope and image processing, while out-of-plane amplitudes were measured with a laser doppler vibrometer. In order to extrapolate the motion of the entire proof mass, the motion of four locations across the proof mass were measured: one point at each of the four corners, where the highest out-of-plane amplitudes would be detected. These locations are shown in Figure 5.18.

The amplitudes of motion were measured by actuating the device in-plane using the on-chip capacitive electrodes, while sweeping the frequency across the natural torsional resonance of the device, which for this resonator was 520 \( Hz \). A combination of 5 \( V_{AC} \)
and a variable DC voltage from $10 - 30 \text{ V}$ was used for actuation. The results of this testing showed near identical in-plane and out-of-plane displacements at each of the four locations, as shown in Figure 5.19; however, a phase delay of the out-of-plane measurements was observed between the locations. This phase delay represents a time delay of the measured response with respect to that of the actuation signal, and varied depending on the location: Two of the data points was $180^\circ$ shifted with respect to the other two data points. These phase delay pairing are illustrated in Figure 5.18, and corresponds to the second parasitic mode shape displayed in Figure 5.17 at 2.6 $kHz$. Based on the location of the capacitive electrodes, this out-of-plane displacement is due to electrostatic levitation as a result of the platform actuation, and should therefore be even further reduced when lower actuation signals are used once the device is placed in vacuum. Even in air, the ratio of the out-of-plane motion to the in-plane motion is only 20 $ppm$. Passing such an small cross-axis rotational rate through an error model of the gyroscope results in a scale factor error of less than 1 $ppb$. 

Figure 5.19: In-plane (left) and out-of-plane (right) motion of the torsional resonator during in-plane actuation. In-plane displacement was measured using optical image processing and out-of-plane displacement was measured using a laser doppler vibrometer.
5.5 Algorithm Implementation

Just as SOI calibration stages were used to obtain preliminary results on the design and calibration potential of the torsional resonator, a macro-scale calibration setup was created for testing the proposed algorithm. This was accomplished by using a commercially available gyroscope and applying a continuous, sinusoidal, high frequency calibration rate that was just barely within the bandwidth of the gyroscope. An artificial drift in scale factor was then induced within the device by applying a large external drift in temperature. The output of the gyroscope was then compared before and after calibration.

5.5.1 Macro-scale Implementation

The macro-scale setup was created from three components: 1) A commercial rate gyroscope, 2) A high frequency electro-magnetic calibration stage, and 3) An external rate table. The rate table provides a theoretically unknown rate, and signifies the inertial rate input from the environment which is being measured. The calibration stage provides a unknown, constant sinusoidal rate for the purpose of compensation. And finally, the gyroscope is used to measure the rate responses. A precision dual-axis rate table from Ideal Aerosmith was used for this testing, along with an in-house fabricated calibration stage. The stage consisted of a torsional platform with a single cental pivot, while linear acceleration was applied tangentially through the use of an electromagnetic motor. The motor was chosen for its low natural frequency of 55 Hz and large working distance of up to 5 mm. Combining these parameters with the 4 inch distance between the application of the force and the pivot, a calibration rate of 150 °/s was anticipated. This rate and bandwidth are well within the parameters of the commercial gyroscope that was used,
which was a model ADXRS649 MEMS gyroscope from Analog Devices. A photograph of the experimental setup can be seen in Figure 5.20, where $\Omega_{IN}$ is the unknown environmental inertial input from the rate table, and $\Omega_{Cal}$ is the known calibration rate from the electromagnetic stage.

### 5.5.2 Normal Operation

Due to the increased mass of the calibration stage, experimental testing revealed that the natural resonance frequency of the electromagnetic actuator reduced from 55 $Hz$ to roughly 10 $Hz$, dropping the calibration rate from 150 $^\circ/s$ to 30 $^\circ/s$. Despite this decrease in the calibration rate, 30 $^\circ/s$ was still adequate for demonstrating the calibration technique. This effect can be seen in Figure 5.21, where the output of gyroscope is observed over time. Within the output, both the 10 $Hz$ calibration rate from the electromagnetic stage, as well as a 0.2 $Hz$ environmental rate from the rate table can be
Figure 5.21: Output of the commercial gyroscope when $\Omega_{IN} = 0.2$ Hz and $\Omega_{Cal} = 10$ Hz. observed as a summation of the two input rates. The 10 Hz calibration signal can then be filtered out from the device output, demodulated to determine the instantaneous scale factor of the device, and used to continually update the output the remaining output of the gyroscope.

It should be noted that there is a necessary assumption that the calibration rate is known with a high degree of precision. Any error within this rate signal directly enters the equation when calculating the instantaneous scale factor, thus influencing the performance of this calibration technique. For this reason, the side-band ratio method is to be employed on the co-fabrication micro-scale device for a continuous measurement of the mechanical amplitude of the platform. In the macro-scale setup, such a method is not possible when using a commercial device, thus restricting the calibration potential.

### 5.5.3 Operation with Induced Scale Factor Drift

With the macro-scale setup operational, the constant environmental and calibration rate inputs were applied to the gyroscope for an extended period of time, as shown in Figure
Figure 5.22: Output of the commercial gyroscope on the macro-scale torsional stage with induced temperature ramp (left), along with the output of the device before and after calibration by the torsional stage (right).

5.21, with rates of 0.2 Hz and 10 Hz, respectively. An external heat gun was then used to elevate the temperature of the gyroscope to its maximum operational temperature of 150 °C, and was then allowed to cool back to room temperature over time. The output and temperature of the gyroscope was monitored over the course of this temperature shift, which is shown in Figure 5.22 (left).

The commercial gyroscope used for this experiment had some degree of internal temperature compensation, which allowed the output of the device to remain relatively constant below temperature of about 50 °C. At higher temperatures, however, the scale factor and bias began to shift considerably. By using the magnitude of the calibration signal to compensate for these drifts, the innate scale factor error of the device could be maintained up to temperatures of 115 °C, during which the raw output drifted by near five times. A comparison of the processed data is shown in Figure 5.22 (right). Because this calibration is through the use of a physical rate signal, not only temperature, any form of scale factor drift can be nulled by this technique.
5.6 Challenges and Future Work

Due to the complexity of the co-fabrication process, operational devices could not be obtained. In this section, a list of the primary challenges of the co-fabrication process are described, along with modifications to the process that could be implemented to obtain improved results.

5.6.1 Challenges

There were a number of challenges that had to be solved throughout the fabrication of the integrated gyroscope and calibration platform, for the process as described. A few of these challenges that were successfully solved included: 1) Etching through the external glass substrate for backside access of the platform during wire bonding, 2) Strong bonding of the small features on the SOI wafer to the external glass substrate, and 3) Characterization of the etch depth during the DRIE etching of the thick wafer stack. However, even though the fabrication process was initially designed to mitigate a number of additional potential challenges, a few of these challenges proved to be insurmountable with the current process.

Four of the unsolved major challenges included: 1) Wafer brittleness due to thin support layers, 2) Retaining wafer cleanliness throughout the multi-mask process, 3) The wire bonding force when making the electrical connections of the gyroscope, and 4) Heat transfer during device etching. Each of these challenges were fundamentally integrated into the successful fabrication and operation of the integrated gyroscope and platform, and adequate solutions could not be found during the fabrication runs. With this new knowledge, another fabrication procedure has been created in an effort to mitigate these
challenges through fundamental changes in the process. These changes are presented below and can be summarized with two core alterations: 1) Internal signal routing, and 2) Complete device encapsulation.

5.6.2 Future Process Flow

The proposed fabrication flow abandons the use of SOI wafers for the creation of the resonators, instead relying on wafer bonding techniques. The reason for this is to embed the routing electrodes between the gyroscope and calibration platform so that the gyroscope signals can be routing off chip, without the need for external electrical connections. Successful implementation of this procedure would alleviate the third unsolved challenge, and enable a solution to the remaining challenges: complete device encapsulation.

By removing the need for external electrical connections of the gyroscope, through etching of the bonded substrate is no longer necessary. In the original process, physical access to the backside of the platform was required to structurally support the platform during wire bonding of the gyroscope; without such support, the generated force would fracture the platform springs. By alleviating the need for internal wire bounds by direct routing of the gyroscope signals, the bonded substrate must now only be partially etched for motional clearance of the platform. This gives rise to multiple benefits. First, by immediately bonding what is essentially a cap wafer over the etched surface, it is immediately protected from further contamination. This same technique can be applied to the device layer, as well, for a fully encapsulated device. By protecting the etched features in this way, resonator cleanliness can be ensured, which solves the second listed challenge. Next, because at least one wafer of the wafer stack remains solid at any given step of
the proposed fabrication, handle wafers are no longer required. This not only aids in the cleanliness of the wafer stack, but also reduces the thickness of the stack during device etching. While this does not completely relieve the heating issue, it does reduce it, which benefits the fourth issue presented. Finally, due to the solid substrate and wafer cap, wafer brittleness should be much less of an issue throughout the fabrication process.

A visual summary of the proposed process is shown in Figure 5.23, where the colors grey, yellow, blue, and purple represent silicon, conductive metal, silicon dioxide, and silicon nitride, respectively. This process requires a number of additional surface mi-
cromachining steps for the fabrication of the embedded electrodes, shown in step one; however, the remaining steps consist of just device etching and wafer bonding.

5.7 Conclusion

Scale factor accuracy of rate gyroscopes is one of the crucial performance metrics for high end applications. As shown in Table 1.4, for inertial grade performance, accuracies of less than 1 ppm are required, which even exceeds the capabilities of the most sophisticated of rate tables. One potential method of overcoming this limitation is through the use of whole angle operation for direct angle measurement; however, an alternative approach has been investigated through the use of an embedded torsional calibration stage. This approach utilizes the stage to transmit a high stability inertial input to the gyroscope for continual scale factor calibration.

To demonstrate the effectiveness of this approach, an error model was developed. It was determined that the calibration potential was directly related to the amplitude and frequency stability of the torsional resonator. To enhance frequency stability, a high Q-factor of the stage was pursued, while amplitude stability could be enhanced algorithmically through the side-band ratio detection method. This method enables a direct measurement of the mechanical amplitude, which could then be used for either feedback or feed-forward calibration.

Prototype calibration stages were created using conventional SOI fabrication, and preliminary rate stabilities were experimentally measured. After passing this stability through the calibration error model, it was determined that scale factor accuracies of less than 20 ppm could be achieved. This was partially enabled by the high Q-factor of
the device of 70,000, which allowed the frequency stability of the device to contribute less than 1 ppm to the scale factor error. While amplitude stability was the limiting factor in this case, these results were obtained prior to the implementation of the side-band ratio method of detection. The parasitic motion of the prototype stages was also measured, which displayed nearly undetectable errors in the transmitted rate signal. Cross-axis rotation on the order of 20 ppm was observed, which was analytically shown to contribute less than 1 ppb to the scale factor error.

In order to qualify the proposed calibration approach, a macro-scale setup was created using a commercially available MEMS gyroscope, an electromagnetic calibration stage, and rate table for environmental input. A dramatic shift in the gyroscope scale factor was then induced by applying a temperature ramp to the device. By utilizing the calibration signal transmitted by the stage, the output of the gyroscope could be compensated to maintain the innate scale factor stability of the device of temperature up to 115 °C. This can be compared to the 50 °C of the sensor normally, which included on-chip thermal calibration. Because a physical rate is being induced, the proposed calibration approach should function for not only drifts due to temperature, but any change in scale factor.

While an integrated micro-scale version of the gyroscope and calibration stage was proposed, the complex co-fabrication process proved to be of a considerable challenge. While many issues were solved throughout the process, a number of unsurmountable problems were encountered that required drastic alterations to the fabrication process. These challenges and alterations are provided, along with a detailed fabrication plan for future implementation.
Chapter 6
Packaging Influences on Device Performance

6.1 Introduction

High-performance MEMS gyroscopes require high Q-factor, small frequency mismatch, and high temperature robustness. These requirements not only provoke many challenges in the design, fabrication, and control of these sensors, but also packaging. Packaging must not only provide a high-level of vacuum to maximize the Q-factor of the resonant device [134], but also prevent the accumulation of stress, which can influence the drift of the resonant frequency with respect to temperature [135], as well as deflections of the proof mass and gaps of capacitive electrodes. This stress accumulation is typically due to the mismatch of bulk material properties.

Wafer-level vacuum sealing is becoming an increasingly promising option to address these concerns, by using silicon as the single bulk material for the resonant device and vacuum cavity [136]. A major challenge that still remains with this approach, though, is through the maximum achievable level of vacuum. When many material are exposed to ambient conditions, particularly metals, these materials absorb molecules which then
desorb when exposed to vacuum. This process is known as ‘out-gassing’. When cavities are encapsulated under high levels of vacuum, post-sealing out-gassing of the encapsulated surfaces typically limit the pressure in the cavity to between 1 and 10 Torr [137], [138]. Through the elimination of encapsulated metal and careful process optimization, it has been shown that this pressure level can be reduced to 5 mTorr [139]; however, there is another mechanism that limits the effectiveness of this sealing technique: hydrogen diffusion.

Hydrogen is known to be able to permeate silicon, preventing the enclosed cavity from reaching a pressure below that of the hydrogen vapor pressure of the ambient environment. In Earth’s ambient atmosphere, this is roughly 0.5 mTorr [139]. While this pressure level may be adequate for a variety of applications, it may still limit the Q-factor of certain resonator designs due to viscous damping. In order to break this theoretical limit, one existing method is to use packaging materials that are more robust to hydrogen diffusion, and encapsulate additional materials which are designed to continue to absorb ambient air molecules after sealing [11]. These gas-absorbing materials are known as ‘getters’ and have been shown to be able to stabilize the cavity pressure [140] and Q-factor of resonant devices [141], with leak rates as low as 10 mTorr per year [140], depending on sealing conditions. The use of this method has enabled resonant Q-factors in excess of 1 Million [75].

Vacuum-sealing with getter material is an involved packaging process, requiring the correct materials, sensor design, die attachment, and sealing process. In this chapter, these parameters are discussed, along with the potential tradeoff that exist between
the achievable Q-factor, attachment strength, and thermal drifts in frequency, modal mismatch, and capacitance.

6.2 Packaging Requirements

There are five major figures of merit when choosing a packaging method: 1) Attachment strength, 2) Attachment rigidity, 3) The degree of stress isolation, 4) The ability to hold vacuum, and 5) Cost. Due to the high variability of cost to many non-technical parameters, this factor will be neglected for now, as the other four elements are discussed below.

6.2.1 Attachment Strength and Rigidity

The strength and rigidity of an attachment process rely on similar parameters, though with very different influences. Both are dependant upon the Young’s modulus of the material interface, yet strength also relies on geometry. While in truth, both factors are interrelated, rigidity has a larger influence on the dynamics of the MEMS device, while strength determines the failure conditions of the attachment.

When an attachment with a low Young’s modulus is used, an additional degree of freedom may appear between the ceramic package and the silicon die. For high-Q resonators, this degree of freedom can lead to an increase in energy loss through the substrate of the device, thus reducing its Q-factor [210]. However, this additional degree of freedom may also lead to an increased robustness to shock, thus improving the survivability of the device to extreme conditions.

When considering the mechanical failure of a packaged device, there are three loca-
tions this failure may occur: 1) The packaging, 2) The resonant structure, and 3) The interface between the structure and the packaging. Failure of the packaging is typically not a concern due to its independent, traditional fabrication methods. However, as the rigidity of the attachment increases, there is less attenuation of a external shock to the MEMS resonator, thus increasing the likelihood of damage. Assuming the resonator can survive the shock, there is also the possibility of a premature interface failure between the device and package. Along with the Young’s modulus of the attachment material, the geometry of the die attachment is an additional influencing factor.

6.2.2 Stress Isolation

In addition to influences on the Q-factor of the device and failure conditions, die attachment may also influence frequency by imparting stress and influencing temperature sensitivity [135]. There are four factors that influence this effect: 1) The temperature at which attachment occurs, 2) The mismatch in the coefficient of thermal expansion between the package and device, 3) The geometry of the die attachment, and 4) The geometry of the resonant structure. This effect will be discussed in greater detail throughout this chapter.

6.2.3 Vacuum Sealing

Most MEMS devices require, at minimum, hermetic sealing to protect their small gaps from environmental debris and contamination. In addition to this, vacuum sealing is also the first requirement to achieve any level of enhanced Q-factor, as shown in Chapter 3. If vacuum sealing is desired, an additional concern is that of maintaining the desired level of vacuum through the life of the sensor.
There are two major factors that may degrade vacuum over time: 1) Leaks between the enclosed cavity and external environment, and 2) Out-gassing of enclosed materials. Most leaks in the package are immediately observable after the attempted vacuum sealing; however, it is possible to encounter small leaks that takes days or even weeks to degrade the enclosed vacuum to an observable level. These effects can generally be eliminated with proper process development; however, leaks may also be present in the form of diffusion. Depending on the material chosen as the barrier between the enclosed cavity and the environment, gasses may eventually diffuse through the material to degrade the enclosed vacuum. Due to its small molecular size, hydrogen is considerably well adapt at this [139], though there are methods to contain it. Luckily, assuming complete hydrogen diffusivity of the cavity, the ambient vapor pressure of hydrogen in Earth’s atmosphere would still result in an enclosed vacuum on the level of 0.5 mTorr. Material out-gassing, however, can degrade vacuum much faster.

Some materials allow molecules to desorb from their surface over time. Organic polymers are notorious for this effect, and is the cause of odors when working with most household adhesive. When such materials are exposed to vacuum, this long-term molecule desorption can quickly degrade any level of vacuum. To a lesser extent, metals behave in a similar way. When exposed to atmosphere, metals typically absorb ambient gasses, which it can then give off over time under vacuum, degrading the vacuum level. If a high, stable level of vacuum is desired within the packaged cavity, only vacuum-safe materials must be used within, including the material used for die attachment.
6.3 Influences of Die Attachment

As indicated in the previous section, die attachment can influence multiple sensor parameters. A few of these influences include: 1) Q-factor in the form of both viscous and substrate losses, 2) Reduction of the operational temperature range, 3) Thermal drifts of frequency and capacitance, and 4) Frequency mismatch due to asymmetric stress. These factors will be discussed below.

6.3.1 Quality factor

As previously mentioned, device Quality factor is both sensitive to die attachment rigidity in the form of substrate losses, as well as material out-gassing in the form of viscous damping. When attempting to maximize device Quality factor, this requires a rigid connection between the device and package, through the use of low out-gassing materials. Fluxless solder is an excellent contender for this and typically requires a metal-to-metal connection. However, as described in the prior section, metal is known to adsorb molecules in atmosphere. Not only can this lead to out-gassing after sealing, but also result in a poor attachment quality due to difficulty in material bonding.

One common method to aid this issue is generally refereed to as a “bake out”, or exposing the material to an elevated temperature prior to attachment to desorb the bonded molecules. Without the use of these effect, attachment generally fails. However, upon x-ray examination of the die attachment area after bake-out, despite the successful attachment, only a small portion of the full area can be seen to result in bonding, Figure 6.1 (left). The dark area in the center of the die represents a strong bond, while the surrounded grey areas represent no attachment. In order to improve this, both the bake out
Figure 6.1: X-ray images of two devices with opposing packaging processes. Both underwent the same temperature profile, but the left was performed in air, while the right was performed in vacuum.

and die attachment procedure must be performed in vacuum. Repeating this procedure with these modifications results in Figure 6.1 (right); full attachment is observed.

There are two potential explanations for this improvement: 1) Completing the bake out procedure in vacuum aids in the desorption of molecules from the metal films, and 2) Without the force of air pressure impeding the solder, it is free to flow across the base of the die. Both of these factors are likely true.

A uniform, void-free attachment area is not only desirable for repeatability, but also to prevent large reductions in vacuum over time. As the device ages, encapsulated air may escape from these voids, resulting in sudden reductions in Quality factor.

6.3.2 Temperature Range and Capacitive Drifts

Depending on the materials of the device and package, the thermal material properties may be different. In particular, the coefficient of thermal expansion is critical for rigid die attachments, as variations in this parameter between the two materials induces interfacial
stress when temperature drifts are experienced. This stress can directly influence the characteristics of the resonator in the form frequency shifts, but also induce deformation in the two materials, as a function of their geometry.

The silicon die is unusually susceptible to this kind of deformation due to its relative size, suspended device layer, and narrow features. As the silicon substrate deforms, this stress directly transfers to the device layer through the anchors, but only weakly through the etched springs to the suspended structure. Depending on the design of the springs and orientation to the imparted stress, deflections can occur, potentially creating unintended mechanical contact of certain features. Should this occur, the resonance properties of the device could either drastically change or simply not function, leading to a potential limiting factor for the maximum operable temperature.

This deformation not only influences the mechanical properties of the device, but also the electrical. In-plane anchor deflection influences the gap of electrostatic electrodes, resulting in capacitive drifts, especially for electrodes of the parallel plate type. Due to the sensitivity of these electrodes, this effect can be difficult to calibrate out of the system due to the lack of on-chip thermometers with the necessary level of precision.

6.3.3 Thermal Coefficient of Frequency and Modal Mismatch

While deflection is a secondary effect on device performance, stress can also directly influence the resonance frequency of the structure. Depending on the attachment size and material properties of the package, the conversion of temperature drift to imparted device stress can change. As such, this effect can influence the thermal coefficient of frequency of the device.
Frequency drifts can have a substantial effect on sensor performance [142]. As shown in Table A.2, the native thermal coefficient of frequency of silicon is about $-30 \text{ ppm/K}$, due to a thermally sensitive Young’s modulus. Much work has been done in order to reduce the natural frequency drifts of silicon, including such methods as utilizing composite resonators [143], [144], as well as temperature-dependant axial force [145]. However, even when no mechanical compensation method is employed, the native value is rarely observed in literature due to packaging effects. Because packaging stress is sensitive to attachment size, die attachment symmetry therefore becomes a concern. This is especially true to high temperature attachment processes, where a large thermal drift is present even at room temperature, thus exacerbating this effect.

For degenerate mode gyroscopes, die attachment asymmetry can therefore result in large frequency mismatches at room temperature, even for mechanical elements of perfect symmetry. Furthermore, the modal mismatch is also a function of temperature, making it difficult to maintain a mode-matched condition, even with electrical tuning. Similar to the capacitive gaps of the electrostatic electrodes, this effect can be compensated using a form of active mismatch control, but precision can be an issue.

### 6.4 Die Attachment

Die attachment is a required step for robust electrical interrogation of the fabricated die. This is true for devices encapsulated at the wafer-level, or exposed after fabrication. In this section, a brief overview of the available methods is provided, along with the employed procedure for in-house device testing and vacuum-sealing, both with or without getter material.
6.4.1 Epoxy versus Solder

While many bonding materials exist, there are two main options: room-temperature organic adhesives, or high-temperature solder. Organic adhesives are commonly used for low-performance applications and experience considerable out-gassing when exposed to vacuum, limiting Q-factor of exposed dies. However, this may not be a concern for encapsulated devices. These types of attachment are also typically soft, granting an additional degree of freedom between the die and package. Depending on the design and frequency of the resonator, it is possible this could reduce the Q-factor through substrate losses; however, it has the benefit of potentially increasing the robustness to shock.

In comparison, fluxless solders exist which out-gas little, making them excellent for vacuum applications. There are several different types of solders that exist for this application, each having different properties. When attempting to achieve a high level of vacuum, it is important to use a 'fluxless' type, as conventional solder may degrade the vacuum over time and at high temperatures due to out-gassing of the flux material [137].

For high-temperature applications after die attachment, eutectic solders are also desirable. When eutectic bonding materials are in good contact with one another and raised in temperature to at least their eutectic point, the materials diffuse into one another, creating a new component with different properties. 80/20 gold/tin solder has this property at a temperature of 280 °C; however, when used to bond together two gold surfaces, the percent of tin decreases during the diffusion, drastically increasing the melting point of the solder in the future. As an example, at 16% tin, the reflow temperature has already risen to 400 °C [146]. This is a particularly useful property for vacuum sealing
applications with getter material, as will be later discussed. Because this is still a high
temperature process, residual thermal stress can still be an issue for the reasons previ-
ously mentioned, but can potentially be minimized through design. In this work, fluxless
eutectic 80/20 gold/tin solder is used for both die attachment and vacuum sealing.

6.4.2 Die Attachment Procedure

There are two main steps for a successful gold/tin solder bond: 1) Baking-out the the
materials to be bonded, and 2) Placing the materials into good contact and reflowing the
solder. It is advisable that both steps are performed under vacuum, requiring the use of a
vacuum furnace, as shown in Figure 6.1. In addition, immediately prior to processing, it
is recommended to clean the empty chamber itself to ensure a clean working environment.
This is especially critical when exposed dies are being packaged. Chamber cleaning is
completed by exposing the chamber to vacuum and elevating the internal temperature
to its maximum level for about an hour before cooling down and wiping the interior with
isopropanol. This also serves as a useful test run for the equipment to identify errors
prior to the involvement of critical materials.

Once the chamber has been cleaned, the packages and silicon dies are then placed into
the chamber, each with their respective die attachment surfaces completely exposed to
the vacuum. Please note that the silicon dies must have gold deposited on the backside
for a strong attachment, for this allows the diffusion of tin, thus completing the bond.
The chamber is then pumped down to vacuum and the temperature elevated to purge
contaminates from the attachment surfaces. To avoid unnecessary aging of the packages,
the back-out temperature is restricted to 320 °C for 2 hours.
Figure 6.2: Temperature profile of a device during die attachment when following the outlined procedure (right), along with photograph of the experimental setup (left).

An 80/20 gold/tin solder preform is then manually placed into the package cavity, along with a graphite frame. The silicon die is then placed within the frame, along with a graphite lid on top of the device, each component with the proper tolerating so that the lid only comes into contact with the perimeter of the silicon die. A weight is then placed on top of the lid, to create good contact of the die and package with the solder. The chamber is then pumped to vacuum and programmed with the temperature profile shown in Figure 6.2. To allow room for the electrical pins of the package, the packages are placed on graphite beams, so the temperature at the solder is slightly less than the programmed temperature of the furnace. This temperature was measured using a bimetallic gauge, which is also shown in Figure 6.2, along with a photograph of the setup. For additional information on the procedure, please see Appendix B.
6.5 Identification of Packaging Stress

In order to experimentally identify the influence of packaging stress on the enclosed mechanical resonator, devices were characterized before and after packaging. Phase 2 quadruple mass gyroscopes were used in this experiment, and were electrostatically actuated and detected before packaging through the use of a microscope-aligned probe card. The devices were then packaged using the previously described eutectic die attachment process, wire bonded, and once again electrostatically characterized. X-ray inspection was also used to identify the final die attachment area after packaging.

A finite element model of this phenomena was also developed and compared to the experimental data. By qualifying the model in this way, the modeling results were then extrapolated for a variety of die attachment geometries and orientations. The results show a significant dependance on the geometry of the resonator, specifically the locations the suspended structure is attached to the substrate. When considering the quadruple mass gyroscope, a small, centrally located attachment point is shown to minimize the influence of die attachment stress. These results also reveal a potential method of thermal compensation, which is also discussed.

6.5.1 Experimental Procedure

For this experimental study, an existing, high-performance sensor design was used: a quadruple mass gyroscope (QMG) [134]. As discussed in Chapter 3, the specific QMGs used were of a Phase 2 design. Each device was X-Y symmetric, consisted of four resonant proof masses, and each mass was coupled to one another through levering structures designed to force an ideal anti-phase motion [13]. The silicon dies were created using
traditional in-house silicon-on-insulator fabrication, as discussed in Appendix B; however, after release, but before packaging was completed, the devices were electrostatically characterized.

This characterization was accomplished using a microscope-aligned probe card, which was specifically fabricated for the Phase 2 QMG electrode locations, as shown in Figure 6.3 (left). Due to the large size of the QMG ($8 \times 8$ mm) and widely spaced electrodes, the probe card was necessary in order to access enough bond pads for differential actuation and detection of both axes. Differential operation was critical for reducing parasitics as the signals traveled along the long, unshielded wires of the probe card. In order to reduce the width of the probed electrodes, as well as characterize any asymmetry that was present in the base device, only a single mass was actuated and detected, as discussed in Chapter 2.

Once the response of the devices was recorded, they were then packaged into symmetric 44-pin leadless chip carrier (LCC) packages, using the previously discussed eutectic die attachment procedure, Figure 6.3 (center). The structures were then wire bonded, placed into a socket, and electrostatically characterized through the package, Figure 6.3
Figure 6.4: Example frequency sweeps of a device before and after packaging. A slightly higher frequency can be observed after packaging.

(right). The identical electrodes that were used for the initial characterization were also used for the second. These two data sets were then compared with one another to observe any changes in the measured response. An example of this comparison can be seen in Figure 6.4, which clearly shows an increase in frequency of the anti-phase mode shape at approximately 3.1 kHz.

During the die attachment procedure, it is know that there is some degree of asymmetry in the solder reflow process as a result of the high temperature process. In order to observe the influence of this effect for each individual device, x-ray inspection was completed. While this procedure is explained in greater detail in Appendix B, an example of the recorded images is shown in Figure 6.5, which clearly identifies the varying degrees of contrast in the image as belonging to the die, solder reflow, and package. In order to quantify the dimensions of the final die attachment area, each attachment area was modeled as an ellipsoid, with principal axes aligned to the silicon die. These measured
Figure 6.5: Packaging procedure and X-ray image of die attachment area. The length and width of the area was measured and compared to the shift in resonance frequency.

Figure 6.6: X-ray images of devices after die attachment, showing the random reflow of solder during attachment.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Silicon</th>
<th>Kyocera A440</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Density:</td>
<td>2330</td>
<td>3800</td>
<td>kg/m³</td>
</tr>
<tr>
<td>Young’s Modulus</td>
<td>170</td>
<td>310</td>
<td>GPa</td>
</tr>
<tr>
<td>Poisson’s Ratio:</td>
<td>0.28</td>
<td>0.23</td>
<td>-</td>
</tr>
<tr>
<td>Coefficient of Thermal Expansion:</td>
<td>2.6</td>
<td>7.1</td>
<td>ppm/K</td>
</tr>
</tbody>
</table>

Table 6.1: Material properties of silicon and Kyocera A440 [148].

dimensions are also shown in Figure 6.5 with variables X and Y. The final dimensions of the die attachment area is truly governed by a stochastic process, with additional examples of the solder reflow seen in Figure 6.6

6.5.2 Modeled versus Experimental Results

In order to form a predictive model of the observed behavior, a 3-D Finite Element Model was created consisting of one quarter of the device and package, Figure 6.7. Symmetry of both the resonant structure and packaging allowed for the reduction of the model to only one of the four proof masses with appropriate symmetric and anti-symmetric boundary condition. The modeled layers included the resonant device layer, silicon substrate, and ceramic package, along with the interfaces between each [147]. The modeling consisted of two stages: 1) calculate the internal stress due to the thermal mismatch between the silicon and ceramic when temperature is decreased from 280°C to 25°C, and 2) determine the anti-phase resonant frequency as a result of this additional stress. These final resonant frequencies were then compared to the initial frequency of the device, before bonding to the substrate at elevated temperatures. The relevant material properties used for both the silicon device and packaging are given in Table 6.1, which has been gathered from Table A.2, Equation A.2, and the package manufacturer [148]. This core model allowed for the manipulation of the bonding area between the silicon substrate and packaging, for
Figure 6.7: Finite element model of the packaging process. The silicon device, along with the ceramic package, was modeled with variable die attachment area. The anti-phase resonance of the device was then measured before and after a large static temperature shift. Due to symmetry of the device and resonance, only one quarter of the device was modeled.

the purpose of observing the effect on resonant frequency. Using this model, two different forms of symmetric die attachment were analyzed: a central, circular attachment area, and frame around the edge of the device, both of varying width.

As the diameter of the die attachment was adjusted in the model, a clear trend in the frequency shift was observed. The results of the modeling is shown in Figure 6.8, along with the experimental data. Despite the low number of data points, a clear trend can be observed in the experimental data, which matches that of the modeled results. There is some discrepancy due to variations in the modeling assumptions and experimental procedure; however, the trend remains consistent: small die attachment diameters lead to reduced shifts in frequency. It is believed that the primary factor influencing this
Figure 6.8: Modeled and experimental results showing the resonance frequency shift versus die attachment diameter.

discrepancy is the translational displacement of the experimental die attachment area, which can be observed in Figure 6.6. To model this effect, the full die would have to be considered, which would quadruple the number of mesh elements. Because the size of the model was already quite considerable, this was not possible with the available computational resources.

The second important result of this modeling is that the influence of die attachment stress was found to be not linearly dependent upon the attachment area. While it may be linearly dependent within the range of the experimental results, as this is extrapolated within the model, clear nonlinearity points can be seen at a number of critical die attachment diameters. When plotting these critical diameters across the surface of the resonator, which can be observed in Figure 6.9, a clear trend appears: each critical diameter corresponds to a crossing of a location where the resonant proof masses attaches
Figure 6.9: Critical die attachment diameters observed in the model of frequency shift before and after die attachment. Each corresponds to when the die attachment reaches a location where the resonant device is anchored to the substrate.

For diameters less than 2 mm, there is virtually no influence of die attachment stress on the resonant frequency. This is because the attachment area is confined to within the internal anchors of the resonator. While this attachment still imparts stress to the die, this stress is easily absorbed by the coupling springs between the masses, allowing the springs that anchor each proof mass to the substrate to remain unchanged. As the die attachment area begins to expand to diameters between 2 and 7 mm, additional area beneath the proof masses is compressed. This imparts stress to the anchoring springs on each of the proof masses, until a maximum frequency shift is achieved at a diameter of 7 mm. This occurs when half of each proof mass resides above the die attachment area and the maximum stress gradient in the resonator is achieved. This begins to decrease
as the diameter expands to 10 mm, which reduces the stress gradient across each proof mass, but of course does not reduce the total imparted stress to zero. From 10 to 12 mm, additional die attachment area continues to raise the total stress in the die, though at a decreased rate as compared to the region between 2 and 7 mm. This modeling shows a considerable design dependance on stress isolation in packaged resonators, where shifts in resonance frequency are not only a function of the total stress in the substrate, but of also the gradient across the resonant proof masses. For symmetric structures, such as the QMG, small, centrally located attachment areas have been shown to minimize the imparted shift in resonance frequency.

In addition to the static frequency shifts due to die attachment, the influence of die attachment area on the thermal coefficient of frequency (TCF) was also examined. This was experimentally tested using three different devices with a range of die attachment diameters: 0.7 mm (green), 2.8 mm (red), and 5.6 mm (blue). The frequency of each of
these devices were then measured at temperatures from 40 to 75 °C and plotted in Figure 6.10. By fitting lines to this data, the TCF of each device was found to be $-23 \text{ ppm/K}$, $-7 \text{ ppm/K}$, and $+36 \text{ ppm/K}$, respectively.

As shown in Table A.2, the native TCF of silicon is approximately $-30 \text{ ppm/K}$. Mismatches in the thermal expansion between silicon and the packaging material can not only influence the static frequency shift of the packaged device, but also how the frequency responds to temperature; in other words, the TCF. With the minimal die attachment area of 0.7 mm, the native TCF of silicon is approached in the experimental data with a measured value of $-23 \text{ ppm/K}$. As the die attachment area increases, this value slowly becomes less negative, and even a positive value. Through intelligent design, such a mechanism could be used as cost-effective thermal compensation of frequency, with no additional fabrication or packaging steps.

### 6.5.3 Reduction of Attachment Area

High die attachment stress can not only influence the resonator frequency, but also displace of the proof masses. In case of the QMG, this motion is most apparent in the hairpin springs. These flexure elements are designed with a high compliance in a single axis, while retaining a high stiffness in the other two axes of motion. High stiffness, however, does not equate to zero deflection. When the large forces generated by the die attachment process are applied to the high stiffness axes of the hairpin springs, deflection can still occur. Axial force on such flexures places one of the beams in compression, and the other in tension. Over the length of the spring this difference in strain accumulates, potentially creating a large lateral displacement at the suspended end of the spring. For
full die attachments, such an effect can be seen in the QMG, yet is alleviated by reducing the die attachment area. This is shown in Figure 6.11 and is an excellent visual indicator of the reduction in stress.

While reducing die attachment area appears to be an ideal choice for reducing the influence of packaging stress on resonator performance, there are some additional challenges to this procedure. First, as previously mentioned, small die attachment areas weaken the interface between the package and die, potentially limiting the maximum survivable shock of the device. Second, despite utilizing an adhesion layer of chrome, hydrofluoric acid (HF) can still penetrate and weaken silicon metallization. Should the die attachment size be reduced below a certain critical area, the silicon metallization may weaken or simply peel off during HF exposure. Without such metallization, the strength of the
The smallest metal pads to survive exposure were approximately 2 mm in diameter. Eutectic die attachment is significantly reduced.

In order to determine this critical die attachment area, multiple sizes of silicon metallization were tested under extended periods of time in liquid HF. When 20% liquid HF is used for release of the in-house resonant structures, the submersion time has never exceeded 90 min. As the maximum release time, 90 min was therefore the goal to maintain adhesion, with photographs of the exposed metal pads being taken in 15 min increments. Six different metallization dimensions were tested, each squares with different side lengths: 170 µm, 340 µm, 510 µm, 1020 µm, 1360 µm, and 2780 µm.

The photographs from each 15 min iteration of the smallest three metallization areas are shown in Figure 6.12 (top-left), along with the final 90 min photographs of the largest three areas (bottom-left). Over the course of the exposure, what appears to be water droplets on the surface of the metallization could be seen to grow, shrink, disappear, and

- Various die attachment sizes
- 90 min liquid HF exposure
- Afterwards, attempted to be stripped using Kapton tape
move. The consistency between photographs, and the fact that the dies were thoroughly dried with compressed nitrogen after removal from the liquid HF, leads to the belief that the observed bubbles are beneath the metallization due to HF penetration. At the end of the 90 min exposure, Kapton tape was then used to attempt to strip the metallization as a form of adhesion test. All of the metallization dimensions failed, except for the largest size of 2780 µm. A photograph of the die attachment after Kapton stripping is seen in Figure 6.12 (right). In this figure, it can also be seen that the metallization pealed off only at the edges and at the locations of the bubbles in the 90 min photograph, which also supports the conclusion that the bubbles are in fact beneath the metal pads.

While die attachment sizes larger than 2 mm appear to be able to survive extended exposure to liquid HF, another option is to replace the liquid HF release step with a vapor HF release step. As discussed in Appendix B, vapor HF release eventually supplanted the liquid release method for a number of reasons. A few of these reasons include increased yield by alleviating stiction, reduced processing times, and increased safety and reliability. When using a substrate temperature of 36 °C, etch rate calibration curves for both in-house fabricated devices and external SOI-MUMPS process are shown in Figure 6.13. Because 15 to 20 µm of undercutting is necessary for release of the in house devices, this corresponds to etch times of only 45 to 60 min.

The current in-house vapor HF etching system is explained in greater detail in Appendix B, along with information on both the liquid and vapor release protocols for comparison. However, another benefit to the use of the vapor HF release is that the backsides of the silicon dies are relatively protected from HF exposure. In order to qual-
Figure 6.13: Vapor-phase hydrofluoric acid release of SOI devices: undercut depth versus exposure time of two different devices (right). Measurements were performed using infrared microscope imaging (left).

ify this, a silicon die with a surface hard mask that still remained after etching was placed into the vapor HF machine upside-down and etched for 60 min. The machines uses an electrostatic force to pull the individual dies against a polymer surface to hold the silicon dies in place during etching. While this does not create a perfect seal, it does create a reasonably robust one. A photograph of hard mask of the silicon die after exposure is displayed in Figure 6.14, which shows that silicon dioxide was etched on the edges of the device, yet the center remained in tack. The thickness of the silicon dioxide at the edges and center was then measured using an ellipsometer. The initial silicon dioxide thickness was approximately 1 µm, and after exposure to the vapor HF, the thickness of the film at the center of the die remained at 10038 ± 119 Å, while on the edge of the
Figure 6.14: Die after 45 minutes of vapor HF exposure with front-side protected. Perimeter was exposed to the vapor, but the silicon dioxide in the center remained.

die the film was nearly completely removed with a thickness of $160 \pm 105 \, \text{Å}$. This result shows that, when using the electrostatic chuck, the center of the backside of the silicon die is protected from vapor HF exposure, alleviating any concerns of weakening in the adhesion of the metallization in the center of the die.

### 6.6 Vacuum Sealing

The vacuum furnace can not only be used for void-free die attachment, as previously discussed, but also as Torr-level vacuum sealing. This process does not allow for the sealing of getter material within the cavity for enhanced vacuum; however, it can still significantly raise Q-factor. Q-factors nearing 1000 have been obtained for QMGs with this process, which is over 50 times improvement compared to the Q-factor values of 20 that have been achieved in air. In this section, this process is described, as well as the
6.6.1 Vacuum Sealing Procedure

The vacuum sealing procedure is nearly identical to that of die attachment, with the exception that the back-out and sealing processes are no longer separate recipes. Immediately before die attachment, the package and device are baked under vacuum for an extended period of time, as described in Figure 6.2. This step is to desorb any molecules from the metal surfaces to prevent any voids in the solder attachment, as shown in Figure 6.1. Preventing voids is even more critical when vacuum sealing devices; however, package aging is also a concern. Exposing the packages to elevated temperatures for extended periods of time can alter the surface, giving the metallization a dull appearance, which is critical process parameters are identified for maximizing the enclosed Q-factor.
believed to be due to diffusion and can influence bond quality. For this reason, there is an upper limit to the temperature and exposure time when implementing such procedures.

Under ideal conditions, devices are to be vacuum sealed shortly after die attachment and preliminary characterization in air. This allows a reduction in the necessary bake-out time and temperature, because less molecules have time to absorb into the surfaces. In order to both minimize the additional bake-out time and temperature, as well as the surface cleanliness prior to sealing, a reduced bake-out process is appended immediately before sealing is conducted, as shown in Figure 6.15 (right).

### 6.6.2 Influence on Q-factor

While the reflow profile of the vacuum sealing recipe remained identical to that of die attachment, several of the bake-out parameters were altered in an effort to maximize the final Q-factor of the device. These parameters included temperature, time, and ramp rate to the reflow temperature. Bake-out temperature and time was shown to have no influence on the final Q-factor, with time periods ranging from 1 to 50 hours, yet the ramp rate from the bake-out temperature to the reflow temperature had a dramatic effect. A plot of the final device Q-factor versus ramp rate is provided in Figure 6.15 (left), which displays the device Q-factor over doubling from 400 to 900 by only manipulating this parameter.

It was initially believed that a low ramp rate would be ideal for creating a uniform temperature distribution during sealing, thus improving process yield. While this may be true, there has not been enough data points to confirm nor deny this hypothesis. It has been confirmed, however, that low ramp rates negatively impact the device Q-factor,
Figure 6.16: Frequency sweeps of a QMG packaged with the temperature profile shown in Figure 6.15. Under vacuum, the Q-factor improves from 20 to 1000 in both modes of resonance.

which is a critical disadvantage. It is believed that this is due to the non-uniform pressures within the vacuum chamber. According to the ideal gas law, heat causes gases to expand. It is also known that convection can not occur in vacuum, leaving only heat transfer by conduction and radiation. For high ramp rates, the hot plate at the base of the vacuum furnace quickly heats to the solder melting temperature. This thermal energy quickly conducts through the packages and heats any residual air within the enclosed packaged cavity. Before the solder melts, the residual air increases in temperature and expands in size, forcing its way into the cooler, lower pressure air of the vacuum chamber. The solder then melts, trapping less air molecules within the cavity, which contributes to a lower overall pressure once the devices cools to room temperature. For slow ramp rates, a differential in the air temperature is never established, which traps the nominal vacuum level of the chamber within the package.
After sealing with this elevated ramp rate, the improvement in device Q-factor can be seen in Figure 6.16, which displays both resonate axes of a QMG before and after sealing. When using Figure 3.30 as a calibration between QMG Q-factor and encapsulated pressure, this corresponds to a final encapsulated pressure of 0.3 Torr, which was improved from a nominal chamber pressure of 1 Torr during vacuum sealing by maximizing the thermal ramp rate.

6.7 Conclusion

There are a number of promising methodologies for device packaging. Wafer level encapsulation appears to be an excellent cost effective method of accomplishing this, which is not only capable of vacuum sealing the device, but also enclosing it completely within silicon to prevent stress accumulation due to thermal mismatches between the device and package. This method, however, is not without challenges. Silicon is a very poor material at preventing hydrogen diffusion, limiting the encapsulated vacuum to approximately 0.5 mTorr, which does not completely alleviate viscous damping of some high-Q resonator designs, such as the QMG. Alternatively, separate packaging can be used.

There are a number of factors to keep in mind when choosing a type of packaging. A few of the primary considerations include: 1) Attachment strength and rigidity, 2) Stress isolation, and 3) Out-gassing concerns for vacuum sealing. These factors have a number of influences on the operational characteristics of the packaged resonator, such as: 1) Q-factor due to viscous and substrate losses, 2) Thermal influences on frequency can capacitive drifts, and 3) Innate frequency mismatch of degenerate devices.

With the goal of maximizing the Q-factor of the resonator for optimized sensitivity,
80/20 Au/Sn eutectic solder has been used for both die attachment and vacuum sealing. It has been shown that such a method can both minimize out-gassing within the package, as well as reduce substrate losses by creating a rigid bond between the silicon die and package. Furthermore, depending on the geometry of the resonator design, a small, central die attachment has been shown to minimize the stress imparted to the resonator from the high temperature attachment process. This has been supported through both experimental data and finite element modeling, and additional guidelines are presented for minimizing stress in other types of resonator designs.

A similar method has also been explored for vacuum sealing of the resonators. Through experimental investigation, it has been determined that the temperature ramp rate from the device bake-out to the solder reflow steps is a significant contributor to the final level of pressure in the enclosed vacuum cavity. By maximizing this ramp rate, the packaged Q-factor has been shown to increase from 400 to 900, which corresponds to encapsulated pressure levels from 1 to 0.3 Torr.
Chapter 7

Conclusion

7.1 Contribution of this Dissertation

Listed below are the primary contributions of this dissertation:

1. **Design and analysis of an anti-phase coupling mechanism for tuning fork resonators** (Chapter 2). This coupling structure was analytically and experimentally shown to selectively stiffen the in-phase resonance of tuning fork resonators, expanding the available design space. Experimentally, this resulted in a modal frequency separation in excess of 120%, which is believed to be the largest in published literature.

2. **Method of quantifying structural asymmetry in multi-mass resonators** (Chapter 2). Due to the design complexity of multi-mass resonators, each discrete component of the design is difficult to quantify, making fabrication asymmetry a challenge to assess. A method is presented of quantifying the discrete mechanical parameters of such systems, independent of capacitive variations. This method can serve as a foundation for electrostatically tuning such systems for improving performance.
3. **Design and fabrication of a quadruple mass gyroscope for reduced acceleration sensitivity** (Chapters 2 and 3). Through the use of the anti-phase coupling mechanism of contribution one, a quadruple mass gyroscope was designed, fabricated, and tested. Due to the enhanced mode ordering, the structure was experimentally shown to have a 20 fold reduction in acceleration sensitivity. Through design, this factor can continue to be improved.

4. **Identification and the development of predictive modeling techniques for the five most critical sources of damping for tuning-fork inertial sensors** (Chapter 3). Within the frequency range of interest, the five primary sources of energy dissipation for mechanical resonators were identified: viscous damping, substrate losses, asymmetry losses, thermoelastic damping, and electronic losses. Each of these mechanisms were modeled and used to design and fabricated a tuning fork structure with a Q-factor of 2.34 million and decay constant of 1300 s, both of which are believed to be the largest in published literature for microfabricated structures.

5. **Control algorithm for the compensation of acceleration sensitivity in Coriolis vibratory gyroscopes** (Chapter 3). Both the rate and quadrature signal of vibratory gyroscopes was experimentally shown to be dependent upon input acceleration through capacitive non-linearities and small changes in electrode gaps. By using the quadrature amplitude to compensate the rate signal, acceleration sensitivity could be effectively removed. This technique could be further improved through PCB modifications, as is proposed in the following section.
6. Derivation of the primary source of rate random walk in the output of Coriolis vibratory gyroscopes (Chapter 3). By modifying the PID controller of the PLL, frequency instability of a gyroscope was modified and correlated to rate random walk. This was further supported by an analytical expression which relates the white frequency noise of the PLL to a phase error within the sense-mode demodulation.

7. Design and fabrication of a tuning fork, frequency modulated accelerometer (Chapter 4). In order to improve sensor alignment stability within multi-axis IMUs, one approach is to integrate multiple sensors on a single substrate. A frequency modulated accelerometer was designed and fabricated to achieve this goal by nominalizing the packaging requirements of gyroscopes and accelerometers, allowing a high vacuum to be beneficial to both. In addition, it is shown that such an approach can effectively decouple the tradeoff between the mechanical scale factor and bandwidth of accelerometers, and may lead to enhancements in sensor dynamic range.

8. Design and analysis of a differential, tuning fork, frequency modulated accelerometer (Chapter 4). When frequency measurements are used as a sensor output, one common strategy is to employ a differential frequency output for cancelling non-inertial drifts. For silicon devices, this is primarily influenced by temperature. To null these common-mode influences, two accelerometers were included on a single substrate with scale factors of opposite polarity, resulting in a differential output.
9. Design and analysis of the influence of mode ordering on enhancing the mechanical scale factor of the tuning fork, frequency modulated accelerometer (Chapter 4). Through utilization of the anti-phase mechanism of contribution one, a non-linear tuning fork FM accelerometer was designed and fabricated, revealing an improvement in mechanical scale factor by over 20 fold at 0 g.

10. Design and analysis of a Coriolis vibratory gyroscope and co-fabricated torsional stage for scale factor calibration (Chapter 5). Scale factor stability is a difficult requirement for high-performance vibratory gyroscopes. In an effort to attain the required performance for inertial navigation, a fabrication process was developed and analysed to include a co-fabricated micro-scale rate stage beneath the sensor. An error analysis of this approach was completed, along with a macro-scale implementation.

11. Assessment of the influence of packaging stress on MEMS resonators (Chapter 6). Packaging stress was shown to have a potentially large influence on the performance of vibratory sensors, through the manipulation of the resonance frequency of such devices. This influence is determined by the method, materials, and size of the die attachment, with a strong influence on the resonator design.

12. Development of a die attachment procedure for minimizing packaging stress (Chapter 6). Through the reduction of the die attachment area by limiting the size of the solder preform used for bonding, it was shown that the influence of packaging stress could be minimized for the specific design in question: a quadruple
mass gyroscope. The size of the attachment and attenuation of the die attachment stress was directly related to the specific design in question, which was supported by both finite element modeling and experimental results.

13. **Development of a vacuum sealing procedure for 0.3 Torr vacuum sealing without getter material** (Chapter 6). Alleviating viscous damping is the first step towards improving the Quality factor of vibratory inertial sensors. While sealing with getter material is preferred for enhanced vacuum sealing, a getterless approach was developed for initial device testing. Out of a number of potential variables, it was shown that the temperature ramp rate during the sealing step of the process was the most critical feature for minimizing the encapsulated pressure after sealing, thus maximizing the resonant Quality factor of the structure.

### 7.2 Future Research Directions

Throughout this work, a number of additional questions have been posed and may lead to future topics of research. These topics are listed below:

1. **Comparative analysis of the various control strategies of vibratory accelerometers and gyroscopes.** There are a number of control strategies for vibratory inertial sensors, which includes open-loop, closed-loop, digital closed-loop, frequency modulation, and for gyroscopes, whole angle operation. Each of these strategies may be optimal for certain operating regimes and environmental conditions; however, it is currently unclear where those boundaries lie. The trade-offs include: white noise, bandwidth, scale factor linearity, power consumption, and
2. Gyroscope parameter estimation through virtual carousaling for feed-forward output compensation. As discussed in Appendix A, gyroscope whole-angle operation allows the structure to be driven along an arbitrary axis, independent of electrode orientation. These same electronics can also be used to virtually rotate this axis within the device and observe changes in the command signals for amplitude, rate, and quadrature, as well as the PLL frequency. Through the use of Equation A.62 and any nonlinearity associated with the on-chip electrodes, the angles and magnitudes of the principal axes of elasticity and damping can be calculated. Assuming this information does not change during device operation, this information can then be used in a form of feed-forward control for compensating these influences.

3. Algorithm development for long-term compensation of bias and scale factor. The bias and scale factors of sensors are known to drift over time. Minimizing these drifts is critical for ensuring an accurate comparison of data points between recalibrations. Depending on the application, sometimes devices are never recalibrated, making long-term drift compensation even more critical as performance continues to improve. A number of techniques can be used to calibration both scale factor and bias. For scale factor, this includes active frequency mismatch tuning through use of the quadrature signal, improving amplitude stability with non-linear side-band-ratio compensation, in-run calibration with a virtual rate, along with either force-to-rebalance control for improved linearity, or whole angle
operation. Long-term bias calibration techniques include temperature compensa-
tion for linear drifts, reducing the phase error within the rate demodulation, or
incorporating a form of mode reversal.

4. **Investigation of the physical origin of the bias instability floor in the 
output of MEMS inertial sensors.** For a majority of applications, white noise 
dominates the resolution of MEMS inertial sensors. As the output of the device 
is averaged over time for a constant inertial input, resolution of the measurement 
 improves, but only to a certain point. Eventually, either the resolution will begin 
worsen, due to random walk or drift, or plateau, due to a bias instability floor. The 
physical origin of this bias instability floor is still unknown, but its experimental 
characteristics have been well observed across disciplines. In DC electronics, it 
is believed to be due to resistance fluctuations within the discrete components; 
however, how this specifically relates to inertial devices is still unclear.

5. **Development of a single-chip system consisting of two accelerometers and one gyroscope.** Inertial sensors which employ in-plane vibrations are gen-
erally capable of higher performance than devices that use out-of-plane motion. 
This is due to a greater feature complexity when designing in-plane motion, along 
with narrow capacitive gaps, both of which is made possible with high-resolution 
lithography. For any given substrate, in-plane motion can be used to create three 
different inertial sensor designs: $z$-axis gyroscopes, along with $x$- and $y$-axis ac-
celerometers. By fabricating a single silicon die with all three types of sensors, a 
robust, lithographic alignment can be maintained between all three devices, both
over shock and temperature, because the entire structure is fabricated from the same material. By including two accelerators, the orientation of the die can always be tracked within a static gravity field, except for any rotations parallel to the field. Such a device would be useful for precision rate measurements, where a high accuracy alignment of the sensitive axis is required. One such example of this would be gyrocompassing. A prototype of such a device is shown in Figure 7.1.

6. Development of a 6-axis IMU consisting of three single-chip systems and capable of continuous self alignment. As a continuation of the development of the high-performance 3-axis chip, three such chips could be oriented normal to one another in space, as shown in Figure 7.2. Through the use of redundant acceleration detection, as well as assuming that the sensors of each specific silicon die remain in alignment to each other, any misalignment between each chip can be monitored. By monitoring this misalignment, a higher degree of accuracy of the IMU output can be maintained than would otherwise be possible.
Figure 7.2: Proposed approach for creating a 6-axis IMU (right) through the use of multiple 3-axis single-chip systems (left). Redundant accelerometers provide continual compensation for chip alignment.

7. Analysis and demonstration of using a multi-mass vibratory gyroscope for the detection of both rate and acceleration. Through the use of a multi-mass resonator, each of the primary vibratory modes can be simultaneously observed using additional signal processing, as shown in Figure 7.3. Such a detection scheme will impart approximately four times the level of electronic noise; however, the benefit is that a single transducer can be utilized for potentially 3-axes of inertial detection: one $z$-axis of rotation, along with $x$- and $y$-axes of acceleration. This result is identical to that of the requirements proposed in Figures 7.1 and 7.2. This technique functions by utilizing the anti-phase resonances of the device for rate detection, while observing the in-phase displacement for acceleration along both in-plane axes. Depending on the driven displacement of the gyroscope, some form of compensation might have to be used for electrode nonlinearities of the force-axis; however, the sense-axis acceleration should be relatively straightforward. An analysis must be conducted to determine the effectiveness of this technique, as
Figure 7.3: Quadruple mass gyroscope with differential detection electrodes applied to each mass. Signals can be combined to isolate each mode shape.

compared to multiple dedicated transducers for each form of measurement.

8. **Selectively damping undesirable vibratory modes, through the use of electronic damping, for improved vibrational resistance.** Anti-phase vibratory gyroscopes can have a number of parasitic resonances which are sensitive to external stimuli. By creating a high-Q anti-phase resonance, then electronically damping the undesirable parasitic modes in a way similar to Figure 7.3, even further vibrational resistance could be created, while maintaining high performance rate.
detection. An alternative application could be to electronically damp amplitude modulated accelerometers within vacuum, in order to place them on the same substrate as high performance gyroscopes. For either application, structural damping could also be modulated in real time by including variable resistors.

9. **Comparative analysis of current detection for vibratory MEMS between transimpedance and transcapacitance detection.** There are three methods of converting current to voltage, as is required by vibratory electrostatic detection: 1) Passively through the use of a resistor, 2) Actively with transimpedance detection, and 3) Actively with transcapacitance detection. Each method contains certain tradeoffs and a clear comparative analysis of each would be useful for future MEMS design. These tradeoffs include: damping imparted to the vibratory structure, noise imparted to the output signal, power requirements, and characteristics of the available operational amplifiers.

10. **Development of control algorithms for the tuning fork FM accelerometer.** In Chapter 4, a high Q-factor tuning fork accelerometer was shown, but performance was not demonstrated. Even though the dynamic range of the device will still be limited by voltage stability, there is still opportunity for algorithm development. By including a voltage-dependent scale factor, this voltage can now be regulated as a function of input acceleration for expanded linear range. By including this influence with a conventional closed-loop control architecture for the in-phase resonance, there is the potential for large improvements in linear range, while minimizing the power requirements of the PID controller.
11. Development of a high Quality factor, temperature-robust FM accelerometer that is not limited by voltage stability. As was discussed in Chapter 4, the proposed FM accelerometer is still vulnerable to voltage stability constraints due to the applied tuning voltage. In order to alleviate this influence, an alternative design should be investigated through the use of axial stress to induce a shift in frequency. Such designs have been previously investigated [127]; however, performance is still limited by relatively low Q-factors. By applying the same methodologies described in Chapter 3 for Q-factor enhancement to an axial stress FM accelerometer design, there is still potential for improved performance.

12. Fabrication modification for removing asymmetric die attachment stress. As discussed in Chapter 6, asymmetric die stress can lead to temperature-dependent variations in the frequency mismatch of gyroscopes, which can be critical for mode-matched operation. This influence can be minimized by reducing the die attachment area; however, the random solder reflow was shown to bond directly to silicon, eliminating the possibility of simply reducing the area of the backside metallization. An alternative approach is to not only reduce the size of the backside metallization, but also etch the unmetallized silicon to prevent bonding. An amendment to the silicon-on-insulator process of Figure B.1 is provided in Figure 7.4, which includes this additional step. After completing the backside metallization and patterning, the photoresist is left on and the backside of the wafer is etched with DRIE. The etching depth does not have to be deep, only 5 to 20 \( \mu m \), and the total depth is not critical; it must only be deep enough to prevent contact with the solder during die
Figure 7.4: Amendment to the process flow of Figure B.1 to reduce die attachment asymmetry.

...attachment. In this way, only the center silicon plateau will bond to the package and the die attachment area can be lithographically defined. Please note that additional silicon pillars should also be included at the corners of the die for increased stability during die attachment; however, these locations should not be bonded.
Bibliography


253


[133] Photograph credit to Alexander Trusov.


263


Appendix A

Fundamentals of MEMS Inertial Sensors

A.1 Introduction

Inertial sensors have a wide range of applications from consumer electronics to high-performance navigation. As these devices continue to increase in performance, while also reducing in cost, size, weight, and power, additional applications continue to rise, leading to even further increase in demand. To continue this advancing trend, detailed knowledge of the performance metrics of these devices are critical for guiding choices considering design, fabrication, and control methodologies.

In this chapter, a thorough theoretical foundation of MEMS inertial sensors is presented through the use of four sections: resonators, noise metrics, accelerometers, and gyroscopes. The first section is dedicated to electrostatic silicon resonators, on which many vibrational MEMS devices are based. This includes material properties, governing equations, electrostatic transduction, front-end electronic design, as well as drive-loop control. The second section examines the various noise characteristics of inertial sensors from an experimental view and supplies a theoretical explanation for their source in both amplitude and frequency modulated outputs. Finally, the third and fourth sections examine the governing equations of both accelerometers and gyroscopes, along with assessments of various types of control algorithms and strategies.

A.2 Resonators

MEMS resonators have many applications, ranging from switches and signal processing to clocks and sensors. Because the design of such structures can be radically different, they are typically separated into two terms: traditional MEMS and RF MEMS. Traditional MEMS tend to operate in the kHz and possibly MHz frequency spectrums and can be focused on a variety of applications, including inertial sensing. RF MEMS operate in the high-MHz to GHz frequency spectrums and are focused on maximizing frequency and stability for signal processing or timing applications.

Due to the high frequencies of RF MEMS, these devices typically consist of a single
beam or bulk-mode resonance and are not well suited for inertial sensing due to their high frequencies and low displacements. For this reason, they will not be discussed in this text. Instead, the focus of this work will be upon devices operating within the Very and Ultra Low Frequency bands (0.3 to 30 kHz), borrowing radio spectrum terminology.

A.2.1 Material Properties of Silicon

Monocrystalline silicon is the single most widely used material in MEMS fabrication. As a semiconductor, it has been extensively used for fabricating the transistors used in integrated circuits (IC); however, between its ability to form thin films with various material properties, as well as having a Young’s modulus that approaches the value of steel, its mechanical properties can be quite useful as well.

A short list of the benefits of using silicon for mechanical resonators include: can be doped (selectively or bulk) to become an electrical conductor or insulator, has a high value of Young’s modulus, is compatible with IC fabrication, can be fabricated on the same substrate as IC read-out electronics, can grow a thermal oxidation layer with has opposing chemical sensitivities when compared to bulk silicon, crystalline planes with various material properties allow precise liquid bulk etch, and anisotropic dry etching technologies exist for high aspect ratio etching.

There are, however, also a number of challenges for this material as well. Some of which include: an anisotropic stiffness matrix, a high thermal sensitivity of the Young’s modulus, dry etching techniques that are influenced by temperature, depth of etch, and conductivity, as well as a brittleness, making resonators prone to fracture.

Because this material is used for fabricating resonant structures, identifying the Young’s modulus of silicon is a critical element. The crystalline structure of silicon has a large effect on how the material behaves to stress, allowing the Young’s modulus to range anywhere from 130-189 GPa depending on the direction the stress is applied. This crystalline structure is preserved during manufacturing, which is made possible by the Czochralski process. During this fabrication, a large cylindrical ingot of silicon is grown from a seed crystal, the orientation of which dictates the final orientation of the boule. The boule is then grounded flat along one or two of its sides. A primary flat, with a length of approximately 30 to 35 mm, is ground parallel to one of the {110} planes, while a secondary flat of approximately 16 to 20 mm is ground at an angle in relation to the primary flat to indicate the wafer orientation and doping: 90° for P-Type {100} wafers, 180° for N-Type {100} wafers, no secondary flat for P-Type {111} wafers, and 45° for N-Type {111} wafers. The boules are then diced into circular wafers and polished to precise thicknesses.

Monocrystalline silicon has a diamond cubic crystal structure, the planes of which can be distinguished using Miller indices, as shown in Figure A.1. In this type of structure, there are three major families of planes, each with different properties: {100}, {110}, and {111}. The various orientations of these planes leads to the anisotropy of silicon’s elasticity; however, certain orientations of the crystal can collapse the stiffness matrix into simpler forms.

A common form of the stiffness matrix aligns the coordinate frame to the cubic structure of the material along the [100], [010], and [001] directions, resulting in an
orthotropic orientation. Because each axis is normal to a \{100\} plane and there is an equivalency in the shear conditions \cite{149}, the stiffness matrix simplifies to Equation A.1 with the following engineering constants: \(E_x = E_y = E_z = 130\text{GPa}, \nu_{yz} = \nu_{zx} = \nu_{xy} = 0.278\), and \(G_{yz} = G_{zx} = G_{xy} = 79.6\text{GPa}\).

While this form is convenient for analytical purposes, standard \(<100>\) silicon wafers have a flat that is aligned to the \{110\} plane from which features are generally aligned. When the \(x\)-axis is defined as normal to the wafer flat, this rotates the stiffness matrix 45\(^\circ\) around the \(z\)-axis, which is normal to the wafer surface. The crystalline coordinate frame becomes: \{110\}, \{-110\}, and \{001\}. Because there are at least two axes aligned to \{110\} planes, this orientation is considered orthotropic \cite{150}, still resulting in a simplified stiffness matrix, which is shown in Equation A.2. From this matrix, the following engineering constants can be calculated: \(E_x = E_y = 169.0\text{GPa}, E_z = 130.0\text{GPa}, \nu_{yz} = \nu_{zx} = 0.278, \nu_{xy} = 0.062, G_{yz} = G_{zx} = 79.6\text{GPa}, \) and \(G_{xy} = 50.9\text{GPa}\).
$$
\begin{bmatrix}
\sigma_{xx} \\
\sigma_{yy} \\
\sigma_{zz} \\
\sigma_{yz} \\
\sigma_{zx} \\
\sigma_{xy}
\end{bmatrix}
= 
\begin{bmatrix}
194.4 & 35.2 & 63.9 & 0 & 0 & 0 \\
35.2 & 194.4 & 63.9 & 0 & 0 & 0 \\
63.9 & 63.9 & 165.6 & 0 & 0 & 0 \\
0 & 0 & 0 & 79.6 & 0 & 0 \\
0 & 0 & 0 & 0 & 79.6 & 0 \\
0 & 0 & 0 & 0 & 0 & 50.9
\end{bmatrix}
\begin{bmatrix}
\varepsilon_{xx} \\
\varepsilon_{yy} \\
\varepsilon_{zz} \\
\varepsilon_{yz} \\
\varepsilon_{zx} \\
\varepsilon_{xy}
\end{bmatrix}
$$

(A.2)

A third type of orientation is also becoming popular for MEMS applications: standard \(<111>\) silicon wafers. While the wafer flat is still aligned to a \(\{110\}\) plane, the direction normal to the wafer surface is a \(\{111\}\) plane. Because no two coordinate vectors are aligned to similar planes, yet still aligned to at least one, this orientation is considered monoclinic [150], with a stiffness matrix shown in Equation A.3. This matrix can be used to derive the following engineering constants: \(E_x = E_y = 169.0\text{GPa}, E_z = 187.8\text{GPa}, \nu_{yz} = \nu_{zx} = 0.180, \nu_{xy} = 0.262, G_{yz} = G_{zx} = 57.8\text{GPa},\) and \(G_{xy} = 67.0\text{GPa}\).

$$
\begin{bmatrix}
\sigma_{xx} \\
\sigma_{yy} \\
\sigma_{zz} \\
\sigma_{yz} \\
\sigma_{zx} \\
\sigma_{xy}
\end{bmatrix}
= 
\begin{bmatrix}
194.4 & 54.3 & 44.7 & -13.6 & 0 & 0 \\
54.3 & 194.4 & 44.7 & 13.6 & 0 & 0 \\
44.7 & 44.7 & 203.9 & 0 & 0 & 0 \\
-13.6 & 13.6 & 0 & 60.4 & 0 & 0 \\
0 & 0 & 0 & 60.4 & -13.6 & 0 \\
0 & 0 & 0 & 0 & -13.6 & 70.0
\end{bmatrix}
\begin{bmatrix}
\varepsilon_{xx} \\
\varepsilon_{yy} \\
\varepsilon_{zz} \\
\varepsilon_{yz} \\
\varepsilon_{zx} \\
\varepsilon_{xy}
\end{bmatrix}
$$

(A.3)

While any crystal orientation can be created depending on the seed crystal used during fabrication, \(<100>\) wafers are by far the most common, though \(<111>\) wafers are starting to become easily found commercially. The main benefit of \(<111>\) wafers is that the in-plane stiffness of the material can be considered isotropic, as shown in Figure A.2, which displays the in-plane Young’s modulus, Poison’s ratio and shear modulus for both \(<100>\) and \(<111>\) standard wafers. Depending on the orientation of the featured etched into the wafer, Young’s modulus can change up to 23.1% for \(<100>\) wafers, while \(<111>\) wafers remain unaffected.

In addition to this, \(\{111\}\) planes are highly resistant to potassium hydroxide (KOH), which is not the case for other crystalline planes. Using this properties, \(<111>\) wafers have been used to undercut and release vibratory structures etched into bulk wafers, without the need for a buried oxide layer or pre-etched substrate for structural release.

One potential complication to the prolific use of \(\{111\}\) wafers is adhesion. Due to the density of silicon atoms along this crystalline plane, there is comparatively a low surface energy which negatively impacts the adhesion of thin films. Values for each of the silicon planes is given in Table A.1, showing approximately a 40% reduction in surface energy of \(\{111\}\) planes, when compared to \(\{100\}\) planes.

Film adhesion and surface energy can be correlated using the Young-Dupré equation, shown in Equation A.4.

$$W_{12} = \gamma_1 + \gamma_2 - \gamma_{12}$$

(A.4)
Figure A.2: Comparison of in-plane elastic properties of standard \( <100> \) (blue) and \( <111> \) (red) silicon wafers when rotating around the normal vector: Young’s modulus (left), Poisson’s ratio (center), and shear modulus (right).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>{100}</th>
<th>{110}</th>
<th>{111}</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atomic Density:</td>
<td>6.78</td>
<td>9.59</td>
<td>15.66</td>
<td>( 10^{14}/\text{cm}^2 )</td>
</tr>
<tr>
<td>Spacing:</td>
<td>5.43</td>
<td>3.84</td>
<td>3.13</td>
<td>Å</td>
</tr>
<tr>
<td>Surface Energy:</td>
<td>2130</td>
<td>1510</td>
<td>1230</td>
<td>mJ/m(^2)</td>
</tr>
</tbody>
</table>

Table A.1: Physical properties of the individual crystalline planes of silicon [151], [152], [153].

where \( W_{12} \) is the work required to cleave two materials with surface energies \( \gamma_1 \) and \( \gamma_2 \), and interfacial tension \( \gamma_{12} \).

Additional properties of silicon that can influence the mechanical behavior of the material is shown in Table A.2.

### A.2.2 Spring and Lever Design

Vibratory silicon structures can be created either by using liquid or dry etching techniques. Liquid etching, such as through the use of KOH, can create beams with angled sidewalls aligned to \{111\} crystalline planes. Dry etching techniques carve vertical

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thermal Coefficient of Elasticity:</td>
<td>-60</td>
<td>ppm/K</td>
</tr>
<tr>
<td>Thermal Coefficient of Frequency:</td>
<td>-30</td>
<td>ppm/K</td>
</tr>
<tr>
<td>Coefficient of Thermal Expansion @ 25°C:</td>
<td>2.6</td>
<td>ppm/K</td>
</tr>
<tr>
<td>Thermal Conductivity @ 25°C:</td>
<td>149</td>
<td>W/mK</td>
</tr>
<tr>
<td>Heat Capacity @ 25°C:</td>
<td>705</td>
<td>kg/mK</td>
</tr>
<tr>
<td>Density:</td>
<td>2330</td>
<td>kg/m(^3)</td>
</tr>
</tbody>
</table>

Table A.2: Physical properties of silicon.
trenches which are normal to the wafer surface, and released through selective undercutting or pre-bonding of an etched substrate wafer. Dry etching allows for the creation of complicated silicon structures which are not dependant upon crystal orientation; therefore, it is used extensively in this work.

Dry etching can be thought of as extruding the 2-D design on the surface of the wafer by the depth of the etch. This can be used to produce straight, thin beams with a rectangular cross-section. The stiffness and mode shapes of such beams is dependant upon both the boundary and forcing conditions, and can be calculated using the Euler-Bernoulli equation, which is shown in Equation A.5.

\[ \frac{d^2}{dx^2} \left( EI \frac{d^2 y}{dx^2} \right) = q \]  

(A.5)

where \( x \) is the distance along the beam, \( y \) is the deflection, \( E \) is the Young’s modulus of the material, \( I \) is the area moment of inertia, and \( q \) is the loading conditions.

Figure A.3 shows a case were a uniform, rectangular beam is assumed to have a constant Young’s modulus across its length (which is true for straight silicon beams). When solving Equation A.5 for the clamped-guided boundary conditions \( y(x = 0) = 0, \frac{dy}{dx}(x = 0) = 0, \frac{dy}{dx}(x = L) = 0, \frac{d^2 y}{dx^2}(x = L) = 0 \) and point load \( F \) at the guided end, the solution for the displacement of the beam, \( y \), is shown in Equation A.6.

\[ y = \frac{FL^3}{12EI} = \frac{FL^3}{Etw^3} \]

\[ F = ky = \frac{Etw^3}{L^3} y \]

\[ k = \frac{Etw^3}{L^3} \]  

(A.6)

where \( I = I_y = \frac{1}{12}tw^3 \) for in-plane motion of the spring and \( I_z = \frac{1}{12}tw^3 \) for out-of-plane motion (not shown), and \( k \) is the equivalent stiffness of the structure given the point load \( F \).
Figure A.4: Common springs designs: (a) crab leg, (b) hairpin, (c) double hairpin, (d) mirrored hairpin, and (e) shuttled hairpin.
Figure A.5: A lever design for scaling and possibly inverting displacement.

Figure A.6: A lever design for scaling and transforming displacement orthogonally.
While simple clamped-guided beams can be used in the suspension of MEMS vibratory systems, they are usually supplanted by competing designs, most of which combine a number of clamped-guided beams in interesting ways. A few common spring examples are shown in Figure A.4, both with certain advantages and disadvantages. The trade-offs between each design generally include: stiffness, displacement linearity, required area on the die, torsion resistance, resistance to substrate deformation, and size of surrounding gaps.

Depending on design and placement, springs can be used to either add or restrict certain degrees of freedom by increasing stiffness \( (k > 0) \). This approach, however, is limited. In certain circumstances, motion must be either redirected or scaled when designing a particular mode shapes \( (k \approx 0) \), or possibly even selectively reduced \( (k < 0) \). To accomplish the latter two tasks, levers can be used.

Levers are essentially thick beams that are intended to resist deformation and transfer motion elsewhere on the chip. They are similar to springs, in that under enough stress they will deflect; however, this is designed to only occur at higher, non-operational frequencies. This technique offers a greater degree of design flexibility and can also be used to design springs with negative stiffness values, as will be discussed in later chapters.

Two common lever designs are displayed in Figures A.5 and A.6. The lever in Figure A.5 is designed to linearly scale displacement, which can be positive or negative depending on where the lever is pinned. The lever in Figure A.6 behaves in a similar way, but also redirects the displacement 90° and the scaling must be negative.

While the mass of the springs can generally be ignored when compared to the large proof mass of the structure, the larger thickness of the levers typically requires their mass to be included in the dynamics of the system for accurate modeling.

### A.2.3 Sources of Damping

The classic damped harmonic oscillator is shown in Figure A.7. Mass, \( m \), and stiffness, \( k \), are the minimum number of elements needed to represent an oscillating system, as the coefficients of acceleration and displacement. To represent the transfer of energy to or from the vibratory system, damping, a coefficient of velocity, must be included. When this coefficient, \( c \), is positive, energy is escaping the system. When this coefficient is negative, energy is being pumped into the system. The first line of Equation A.7 shows the classic second order oscillator, with forcing term \( F \).

\[
\begin{align*}
    m\ddot{x} + c\dot{x} + kx &= F \\
    \ddot{x} + \frac{c}{m}\dot{x} + \frac{k}{m}x &= \frac{F}{m} \\
    \ddot{x} + \frac{\omega}{Q}\dot{x} + \omega^2 x &= \frac{F}{m} \\
    \frac{c}{m} = \frac{\omega}{Q} = 2\zeta\omega &= \frac{2}{\tau}
\end{align*}
\]
Depending on the phenomena being evaluated, damping can be expressed by either the damping coefficient, $c$; Q-factor, $Q$; damping ratio, $\zeta$; or dissipation time constant, $\tau$. How each of these terms relate to one another is shown in Equation A.7. Damping ratio is useful when defining the various analytical regimes of motion, with $\zeta = 1$ being defined as critical damping, representing a dissipation level that minimizes the time to reach steady-state after an impulse. Another important note is that each factor along the last line of Equation A.7 represents the half-power bandwidth of the resonator, or the span of angular resonance over which the resonance power is greater than half the value at the resonance peak. In this work, $Q$ and $\tau$ will be used extensively.

When damping is expressed in terms of $\tau$ and the forcing term is set to a sinusoidal forcing condition of $F = F_0 \cos(\omega_d t + \phi_d)$, the solution for the displacement, $x$, can be represented as a linear combination of the homogeneous solution, $x_{\text{homogeneous}}$, and steady state solution $x_{\text{steady state}}$. The solutions to these components are given in Equation A.8 for the underdamped case ($\zeta < 1$).

\[
\ddot{x} + \frac{2}{\tau} \dot{x} + \omega^2 x = \frac{F_0}{m} \cos(\omega_d t + \phi_d)
\]

\[
x = x_{\text{homogeneous}} + x_{\text{steady state}}
\]

\[
x = A_h e^{-t/\tau} \sin(\omega_h t + \phi_h) + A_s \cos(\omega_d t - \phi_s)
\]

\[
\omega_h = \sqrt{\omega^2 - \left(\frac{1}{\tau}\right)^2}
\]

\[
A_s = \frac{F_0/m}{\sqrt{(\omega^2 - \omega_d^2)^2 + \left(\frac{2\omega_d}{\tau}\right)^2}}
\]

\[
\phi_s = \tan^{-1}\left(\frac{2}{\frac{\omega_d}{\tau}} \frac{\omega_d}{\sqrt{\omega^2 - \omega_d^2}} \right) - \phi_d
\]

\[
\phi_h = \tan^{-1}\left(\frac{\omega_h (x_0 - A_h \cos(\phi_s))}{\dot{x}_0 + \frac{1}{\tau}(x_0 - A_s \cos(\phi_s)) - A_s \omega_d \sin(\phi_s)}\right)
\]

\[
A_h = \frac{x_0 - A_s \cos(\phi_s)}{\sin(\phi_h)}
\]

where $x_0$ and $\dot{x}_0$ are the initial position and velocity of the resonant mass.
Placing damping in terms of the dissipation time constant, $\tau$, is useful when the initial conditions dominate the motion of the structure and the forcing conditions are either absent or small. The reason for this is because $\tau$ determines how long the initial conditions will influence the system, with the amplitude of the homogeneous solution decaying as a function of $e^{-t/\tau}$, which approaches zero as time, $t$, goes to infinity.

A graphical representation of the dissipation time constant, $\tau$, is shown in Figure A.8. The figure shows the time-domain response of Equation A.8 when the resonant mass is unforced, $F_0 = 0$, and has normalized initial conditions of $x_0 = 1$ and $\dot{x}_0 = 0$. For graphical purposes, $\omega = 10$ Hz and $\tau = 1$ sec. The dissipation time constant represents the amount of time needed for the amplitude of resonance to decay to a value of $e^{-1}$, or approximately 36.8%. As a note, it takes 7 time constants, $7\tau$ sec, for an amplitude to decay to just below 0.1% of its initial value.

When the forcing conditions dominate the response of the resonant mass by assuming enough time has passed for any initial conditions to decay from influence, damping may be better described by the quality factor, or Q-factor, $Q$, of the resonator. To examine this, consider the same second order oscillator of Equation A.8. Assume enough time has lapsed that any initial conditions can be ignored and instead of forcing the resonant mass directly, the previously grounded support of the spring is now given a sinusoidal displacement of $X_0\cos(\omega_d t)$. This is equivalent to a forcing condition of $\frac{F_0}{m}\cos(\omega_d t + \phi_d) = \omega^2 X_0\cos(\omega_d t)$. Damping is also converted from the form of the dissipation time constant, $\tau$, to a Q-factor, $Q$. The new equation of motion, along with the solution for the amplitude and phase of the proof mass, is given in Equation A.9.
The solution for the magnitude, $A_s$, and phase, $\phi_s$, of the resonant mass in Equation A.9 can then be normalized against the amplitude of the base of the spring, $X_0$, and pi, $\pi$. When plotted against an input frequency, $\omega_d$, which has been normalized against the natural resonance frequency of the system, $\omega$, the results are shown in Figure A.9 for various values of Q-factor.

When actuated at the resonance frequency, $\omega_d = \omega$, Q-factor behaves as the gain of the system, its value serving as a multiplier of how the amplitude of the input motion will be transmitted to the system dynamics. For example, if a single-mass resonator with a Q-factor of 1 million was moved sinusoidally at its resonance frequency with an amplitude of 5 Å, or roughly the distance between individual atoms of silicon according to Table A.1, the resultant motion of the proof mass of the system would be 0.5 mm; a motion distinguishable to the naked eye.

Energy loss arises from a number of sources. When attempting to consider which sources will be dominant for a given oscillator, the most critical property of the oscillator is its resonance frequency. Borrowing radio spectrum terminology, the frequencies of interest in this work comprise the Very and Ultra Low Frequency bands (0.3 to 30 kHz). Within these frequency bands, and for the designs studied, energy loss is dominated
by several loss mechanisms: viscous damping, substrate losses, structural asymmetry, thermoelastic damping (TED), and electronic damping. At higher frequencies, additional loss mechanisms include surface conditions, material defects, and phonon interactions, among others [154].

\[
\frac{1}{Q_{\text{Total}}} = \frac{1}{Q_{\text{Viscous}}} + \frac{1}{Q_{\text{Substrate}}} + \frac{1}{Q_{\text{Asymmetry}}} + \frac{1}{Q_{\text{TED}}} + \frac{1}{Q_{\text{Electronic}}} + \frac{1}{Q_{\text{Other}}} \tag{A.10}
\]

Each loss mechanism can be thought of as a separate Q-factor for the system, or rather, the Q-factor the system would have if no other loss mechanisms were present. Multiple sources of damping can be combined inversely according to Equation A.10. While many of these energy loss mechanisms have been previously discussed in the main text, a brief description of each follows.

Viscous damping: Energy loss due to collisions with air molecules. When thinking of air as a fluid, a vibrating structure must push past molecules to deflect. This can arise due to two mechanisms: squeeze film damping and slide film damping. Squeeze film damping occurs when the vibrating structure is required to push through and deflect air molecules, while slide film occurs when the structure simply slides along a viscous surface. For capacitive MEMS, squeeze film frequency dominates due to the large numbers of parallel plate electrodes. While this mechanism can be removed by evacuating the cavity which contains the resonator through exposure to vacuum, it is typically the greatest source of energy loss for structures exposed to air.

Substrate losses: Energy loss due to externally propagating stress waves through the substrate of the resonator. As a one-dimensional resonator vibrates, the vector of momentum oscillates in magnitude along a single line of motion. This oscillating momentum causes deflection in the substrate, producing stress waves that expand outward from the structure. This form of energy loss can be mitigated through the use of anti-phase resonances; however, a certain amount of energy loss can remain due to any unbalanced components. This residual energy loss is typically inconsequential at low frequencies, but increases in magnitude as the oscillating frequency increases due to design modifications.

Asymmetry losses: Energy loss due to stiffness or mass imbalance in anti-phase resonators. A form of substrate loss, this loss mechanism is only relevant when utilizing anti-phase resonances. Should the anti-phase resonator have any asymmetry in stiffness or mass across the mirrored plane of symmetry, a momentum imbalance occurs which can lead to an increase in energy loss. Asymmetry is always present in the form of fabrication imperfections, so mitigation of these fabrication imperfections may be required, depending on the tolerances of the fabrication method.

Thermoelastic damping (TED): Energy loss due to internal material friction. Any material that deflects has internal stress during deflection. This stress is fundamentally linked to the thermal domain through material’s coefficient of thermal expansion. When considering a deflected beam, one side is in tension while the other is in compression. Through the material’s coefficient of thermal expansion, this raises the temperature at one side of the beam, while decreasing it on the other, creating a thermal gradient. As this thermal gradient flows across the width of the beam, entropy rises resulting in energy loss.
Electronic losses: Energy loss of capacitive devices due to the interface of front-end electronics. These types of losses only appear when the output current of the resonator is converted to voltage. As will be later discussed in this chapter, the velocity of capacitive MEMS resonators can be detected through an output current. Because current can not be read directly, this signal must first be converted to voltage. This conversion requires energy which can either be drawn from the resonator through the use of a passive resistor, thus damping the structure, or externally applied using a transimpedance amplifier. Even through the use of a transimpedance amplifier, delays in the conversion can still result in energy loss of the structure, though this is typically inconsequential.

A.2.4 Electrostatic Actuation

As mentioned previously, two of the benefits to the use of silicon is that: 1) It can be doped during manufacturing to modify its electrical properties, making it either a conductor or an insulator, and 2) Dry etching techniques exist that can etch vertical trenches in the silicon. When a conductive silicon layer is bonded to an insulator and fully etched to the insulator, intricate regions can be fabricated and electrically separated. Depending on design and after a timed etch of the buried insulator where exposed to the surface, certain regions can also be made to move freely. This process can be made to produce variable capacitive gaps between fixed electrode structures and the moveable resonator.

In order to derive the impact of such a structure, capacitance is defined in Equation A.11.

\[ C = \frac{Q}{V} = \frac{Q}{Ed} \]
\[ \frac{Q \varepsilon A}{\rho d} = \frac{\varepsilon A}{d} = \frac{\varepsilon tw}{d} \]

where \( C \) is capacitance, \( Q \) is the charge on the capacitor, \( V \) is the voltage across the capacitor, \( E \) is the electric field between the plates, \( d \) is the distance between the plates, \( \varepsilon \) is the permittivity between the plates (\( 8.854e - 12 \) for both air and vacuum), \( \rho \) is the charge density on the plates, \( A \) is the area that the plates overlap, and \( t/w \) are the length and width of the plate overlap.

For linear mechanical displacement, work can be defined as the integral of force over displacement. An analogous representation in the electrical domain is the integral of voltage over charge. When the work of a capacitor is considered, voltage can be separated into capacitance and charge using Equation A.11, and the integral solved. For convenience, charge can then be converted back to voltage and capacitance in the solution. This process is shown in Equation A.12, showing that the work performed by a capacitor is equal to \( \frac{1}{2}CV^2 \).

\[ U = \int Fdx = \int Vdq = \int \frac{q}{C}dq = \frac{1}{2}Q^2 = \frac{1}{2}VQ = \frac{1}{2}CV^2 \]
\[ F = \frac{d}{dx}U = \frac{d}{dx}\left(\frac{1}{2}CV^2\right) = \frac{1}{2}\frac{dC}{dx}V^2 \]
When considering linear mechanical displacement, force can also be written as the derivative of work over displacement. Inserting the electrical work performed by a capacitor and realizing that only capacitance will change with displacement, the force between the capacitive plates becomes \( \frac{1}{2} \frac{dC}{dx} V^2 \).

Two basic types of capacitive electrodes exist with fundamentally different values of \( \frac{dC}{dx} \): 1) Variable area electrodes, and 2) Variable gap electrodes. The basic behavior of these two different structures are shown in Figure A.10. Variable gap electrodes are relatively intuitive: two flat electrode planes with a variable distance between them. Variable area electrodes, however, require a bit more complicated of a structure. To prevent the plates from pulling together, two techniques are used: 1) Two static plates are placed on either side of a single movable plate, resulting in a cancelation of pull-in force when in equilibrium, and 2) The flexures of the movable plate are oriented to produce only the desired motion. By placing a series of these interdigitated plates together, the capacitive gradient can be increased.

Variable gap electrodes are typically called ‘parallel plate electrodes’, while variable area electrodes are called ‘comb drives’. Figure A.11 displays close photographs of each of these electrode types.

To derive the capacitive gradient of these devices with displacement, \( \frac{dC}{dx} \), the formula for capacitance in Equation A.11 is updated with a displacement variable for each electrode type. For variable area electrodes, width of the plates, \( w \), is replaced with \( w \pm x \) to reflect the changing dimension. For variable gap electrodes, gap between the plates, \( d \), is replaced with \( d \pm y \). The capacitive gradient can then be calculated for each electrode.
type and their respective displacement variables, as shown in Equation A.13.

\[
\left( \frac{dC}{dx} \right)_{\text{Area}} = \frac{d}{dx}\left( \varepsilon t \left( \frac{w \pm x}{d} \right) \right) = \pm \frac{\varepsilon t}{d} \\
\left( \frac{dC}{dy} \right)_{\text{Gap}} = \frac{d}{dy}\left( \varepsilon tw \frac{d \pm y}{d} \right) = \frac{d}{dy}\left( \varepsilon tw(d \pm y)^{-1} \right) = \mp \frac{\varepsilon tw}{(d \pm y)^2}
\]  

(A.13)

Parallel plates electrodes are innately nonlinear for large displacements because the capacitive gradient is still a function of the instantaneous gap between the plates, \( d \pm y \). Their benefit, though, is that they are simple and have a higher sensitivity, considering the width of the plate, \( w \), can be made much larger than the gap, \( d \).

Comb drives have the benefit of observing a linear capacitive gradient with displacement, allowing large, linear displacement at the cost of complexity and sensitivity. This derivation, however, assumes infinitely long comb fingers. When implemented, there is still a nonlinear component due to the edge effects of each plate. This can be made effectively small through design, at the cost of additional electrode space.

Combining the electrostatic force equation in Equation A.12 with the specific capacitive gradient for the electrode design in Equation A.13, an equation for the force applied to the resonant structure can be derived, as a function of the voltage across the capacitor. This voltage is assumed to have both a DC component, \( V_{DC} \), as well as an AC component, \( V_{AC} \sin(\omega t) \). Squaring this voltage produces components at three different frequencies: constant, \( \omega \), and \( 2\omega \). This derivation is shown in Equation A.14.

\[
V = V_{DC} + V_{AC} \sin(\omega t) \\
V^2 = V_{DC}^2 + \frac{1}{2} V_{AC}^2 + 2V_{DC}V_{AC} \sin(\omega t) - \frac{1}{2} V_{AC}^2 \cos(2\omega t) \\
F = \frac{1}{2} \frac{dC}{dx} V^2 \\
F = \frac{1}{2} \frac{dC}{dx} \left( V_{DC}^2 + \frac{1}{2} V_{AC}^2 + 2V_{DC}V_{AC} \sin(\omega t) - \frac{1}{2} V_{AC}^2 \cos(2\omega t) \right)
\]  

(A.14)

For singled-sided actuation, where a single electrode is used to force a resonant structure in one direction, this is the expression for the generated force. For differential actuation, two identical electrodes are used to force a resonant structure, oriented in opposite directions and given opposite polarity AC signals. When this occurs and the combined force on the resonant structure is solved, the DC and \( 2\omega \) forcing components cancel, while the \( \omega \) components combine. This derivation is shown in Equation A.15.
\[ V_{0^\circ} = V_{DC} + V_{AC} \sin(\omega t) \]
\[ V_{180^\circ} = V_{DC} - V_{AC} \sin(\omega t) \]
\[ V_{0^\circ}^2 = V_{DC}^2 + \frac{1}{2} V_{AC}^2 + 2V_{DC}V_{AC} \sin(\omega t) - \frac{1}{2} V_{AC}^2 \cos(2\omega t) \]
\[ V_{180^\circ}^2 = V_{DC}^2 + \frac{1}{2} V_{AC}^2 - 2V_{DC}V_{AC} \sin(\omega t) - \frac{1}{2} V_{AC}^2 \cos(2\omega t) \]  \hspace{1cm} (A.15)
\[ F_{\text{Diff}} = \frac{1}{2} \frac{dC}{dx} V_{0^\circ}^2 - \frac{1}{2} \frac{dC}{dx} V_{180^\circ}^2 \]
\[ F_{\text{Diff}} = 2 \frac{dC}{dx} V_{DC} V_{AC} \sin(\omega t) \]

The results from Equations A.14 and A.15 reflect two different actuation strategies. Differential actuation with both DC and AC components is ideal when a large force is desired, supplied at the same frequency of the AC actuation signal. Single-sided actuation with only an AC component can be used to supply a forcing frequency twice higher than the AC signal, albeit with a DC offset. As will be later discussed, the higher forcing frequency can be a useful advantage to eliminating the parasitic feed-through of the excitation voltage.

### A.2.5 Electrostatic Frequency Tuning

Due to the nonlinearity of parallel plate electrodes, when a DC voltage is present across the gap, it also produces an effect that mimics the behavior of a spring with negative spring constant. This effect is referred to as an electrostatic spring and can be derived through the Taylor series expansion of the electrostatic force, as shown in Equation A.16.

\[ \sum_{i=0}^{\infty} \frac{f^n(a)}{n!} (x - a)^n \]
\[ F = -\frac{1}{2} \frac{\varepsilon tw}{(d + y)^2} V^2 \]
\[ F = -\frac{1}{2} \frac{\varepsilon tw}{d^2} V^2 + \frac{\varepsilon tw}{d^3} V^2 x - \frac{3}{2} \frac{\varepsilon tw}{d^4} V^2 x^2 + \frac{2\varepsilon tw}{d^5} V^2 x^3 + \ldots \]  \hspace{1cm} (A.16)

For small displacements, only the first two terms of the Taylor series expansion can be considered, which can be thought of as a static force plus a spring. The spring constant can then be solved, as shown in Equation A.17.
\[ F = F_0 - kx \]
\[ F \approx -\frac{1}{2} \varepsilon tw \frac{d^2 V^2}{dx} + \varepsilon tw V^2 x \]
\[ F_0 \approx -\frac{1}{2} \varepsilon tw V^2 \]
\[ k \approx -\varepsilon tw \frac{d^3 V^2}{dx^3} \]  

Because the voltage across the gap is squared, this effect can only be used to reduce the stiffness of a structure.

### A.2.6 Capacitive Detection

Electrostatic electrodes can not only be used to force the resonant structure, but also sense its motion. This can be derived through the definition of current, which is the rate of change of electrical charge over time. When this is applied to the charge across a capacitor, Equation A.11 can be used to convert this definition to the rate of change of both capacitance and the voltage across the gap. Expanding this equation results in the linear combination of two terms: one is a function of the rate of change of the gap’s magnitude, which is equivalent to the velocity of the resonant structure, with a second term a function of any change in the bias voltage. This is shown in Equation A.18.

\[ i = \frac{dQ}{dt} = \frac{d}{dt}(CV) = \frac{dC}{dx} \dot{x}V + CV \]  

As derived above, current produced by an electrostatic electrode can be used to observe the velocity of a resonator; however, this requires a voltage bias across the capacitor. There are several biasing schemes that can be used to produce this measurement, which are further complicated when including electrostatic forcing electrodes as well. The trade-offs between each biasing scheme include: the number of electrodes required, the degree of static force applied to the resonator, noise in the output, the degree of parasitic feedthrough of the AC driving signal in the output, and the number of supporting electrical components. In this section, two biasing schemes will be examined to display a few of the primary tradeoffs that must be considered. In the next section, a third biasing scheme will be presented which was used in this work.

The first biasing scheme what will be discussed is the 2-port, single-sided electrode scheme, as shown in Figure A.12. This figure displays the scheme by overlaying the electrical components on a cross-section of a device fabricated with the silicon-on-insulator process which is discussed in Appendix B. In this scheme, there are two electrostatic electrodes, each with a variable capacitance that is used to either force the structure, \( C_D \), or sense the resultant motion, \( C_S \). The proof mass is biased with a static voltage, \( V_C \), with the input and outputs of the device present in the form of an AC voltage, \( V_{IN} \),
Figure A.12: Cross-sectional view of an SOI resonator with single-sided forcing and detection (top), and with an electrical overlay (bottom).

and an output current, \( i_{\text{OUT}} \). In addition, there is a degree of parasitic feed-through directly between the drive and sense electrodes. This is produced by electrical conductivity through the substrate of the device, \( R_s \), along with two parasitic capacitors that separate these electrodes from the substrate, \( C_{P1} \) and \( C_{P2} \). To help reduce these influences, the substrate is typically grounded; however, depending on the resistance of the substrate and placement of the electrodes, residual influences can still be identified.

When examining the circuit diagram of Figure A.12, the total output current, \( i_{\text{OUT}} \), can be seen to be a summation of the current across the variable sense capacitor, \( C_S \), along with the static parasitic capacitor, \( C_{P2} \). This relation can be expanded using Equation A.18 to form Equation A.19.

\[
  i_{\text{Sing}} = i_{\text{OUT}} = i_{C_S} + i_{C_{P2}} \\
  i_{\text{Sing}} = \left( \frac{dC_S}{dx} \dot{x} V_C + C_S \ddot{V}_C \right) + \left( C_{P2} \ddot{V}_{INa} \right) 
\]  \hspace{1cm} (A.19)

Using this scheme, the generated force on the resonant structure will follow Equation A.14, where \( V_{AC} = V_{IN} \) and \( V_{DC} = V_C \). The proof mass bias serves as the DC component for AC plus DC actuation, which still results in forces terms at three different frequencies: constant, \( \omega \), and \( 2\omega \). Assuming the frequency of \( V_{IN} \) is at the resonance frequency of the device, \( \omega \), Equation A.19 has two components at the resonance frequency of the device: the motional current \( \left( \frac{dC_S}{dx} \dot{x} V_C \right) \), along with the parasitic feed-through \( \left( C_{P2} \ddot{V}_{INa} \right) \). Because \( V_C \) is constant, the term \( C_S \ddot{V}_C \) is zero.

The parasitic feed-through is an important consequence of this detection scheme. First, it is at the frequency of device operation, which makes the parasitic signal indistinguishable from the desirable output: the motional current of the device. Second, the
parasitic feed-through may also be variable over time due to amplitude changes of $V_{IN}$ or resistance changes in the substrate, $R_S$. The first effect can influence the observed dynamics of the system. Depending on the relative values of the parasitic and motional capacitance, the parasitic signal has the potential to significantly overwhelm the dynamics of the resonator. This effect can limit the observability of the resonator’s response, with even moderate influences limiting the effectiveness of certain control schemes, such as phase-locked-loops. The second effect can also lead to increased measurement uncertainty by exacerbating noise and drift characteristics.

One method that is capable of reducing these effects while still using this biasing scheme is by setting the frequency of the actuation voltage to be one half of the natural resonance frequency of the device ($V_{IN} = \|V_{IN}\| \sin(\frac{1}{2} \omega t)$). This strategy is still able to actuate the resonator into resonance, while significantly separating the motional and parasitic current at the output in the frequency domain. The parasitic signal can then be removed through filtering. The consequence for this scheme is a reduction in the magnitude of the actuation force.

An alternative to the 2-port, single-sided electrode scheme is the 2-port, differential electrode scheme. A schematic of the differential scheme can be observed in Figure A.13. Similar to Figure A.12, the differential scheme is modified with the inclusion of differential electrodes at both the drive and sense electrodes.

Through the use of the circuit diagram of Figure A.13, the total output current is observed with two differential values, $i_{OUT1}$ and $i_{OUT2}$, separated from each other by 180° of phase. These terms are later subtracted using electronics to form a differential output; it is this output that is now being considered in Equation A.20.
\[ i_{\text{Diff}} = i_{\text{OUT1}} - i_{\text{OUT2}} \]
\[ i_{\text{Diff}} = (i_{C_{s1}} + i_{C_{p3}}) - (i_{C_{s2}} + i_{C_{p4}}) \]
\[ i_{\text{Diff}} = \left( \left( \frac{dC_{s1}}{dx} \dot{x}V_C + C_{s1} \dot{V}_C \right) + (C_{p4} \dot{V}_{INa}) \right) - \left( \left( \frac{dC_{s2}}{dx} (-\dot{x})V_C + C_{s2} \dot{V}_C \right) + (C_{p4} \dot{V}_{INb}) \right) \]
\[ i_{\text{Diff}} \approx 2 \frac{dC}{dx} \dot{x}V_C \]

(A.20)

Using this scheme, the generated force on the resonant structure will follow Equation A.15, canceling the DC and \( 2\omega \) forcing components to produce a clean forcing signal at the resonant frequency of the device, \( \omega \). Equation A.20 also shows the benefit of differential detection: Similar to Equation A.19, the \( C_{s\#} \dot{V}_C \) terms are zero due to the constant biasing voltage. The motional currents, \( \frac{dC_{s\#}}{dx} \dot{x}V_C \), also add together due to their opposite geometric arrangement. But it is the parasitic feed-through that is attempted to be eliminated by this scheme. By placing the differential detection electrodes within close proximity with each other, the hypothesis is that the distance between the drive and sense electrodes are roughly equivalent, resulting in a first-order cancelation of this effect.

These are only two examples of many resonator biasing schemes. As a few examples, these two schemes can be mixed (such as single-sided drive with differential detection), drive and sense the resonator from a single port (by biasing the resonator with the AC driving voltage) [155], or even switching the polarity of the detection electrodes to pull the detection current from the proof mass of the structure [156]. In the next section, an additional biasing scheme will be analyzed, which was utilized in the bulk of this work.

### A.2.7 Electromechanical Amplitude Modulation (EAM)

On the die level, electromechanical amplitude modulation simply involves replacing the static biasing voltage for the capacitive detection electrodes, \( V_C \), with a high frequency AC signal [157]. Assuming this frequency is substantially greater than the vibratory motion of the resonator, the two frequencies become modulated at the capacitive sense electrodes (\( \dot{x}V_C \)). When observing the frequency spectrum of the output, this effect produces sidebands from the carrier signal, separated by plus and minus the motional frequency.

This effect can be seen through Equation A.21 for the 2-port, single-sided electrode scheme, but there are a few important notes to this derivation. First, these equations assume that the biasing scheme of the device still involves an AC plus DC driving voltage. In Figure A.12 this was accomplished by applying an AC signal to \( V_{IN} \) and allowing the voltage to mix across the driving capacitor using the DC signal \( V_C \). Because \( V_C \) has since been replaced with a high frequency sinusoid, the impact of \( V_C \) on the drive force is assumed to be zero; it will create a high frequency force on the resonator, however,
Figure A.14: Spectral density plots demonstrating the electromechanical amplitude modulation technique.

The dynamics of the structure effectively filter the response. In order to continue to drive the structure with an AC plus DC voltage, the mixing of these components must now be preformed off-chip and applied together at $V_{IN}$.

$$V_{C} = A_{C} \sin(\omega_{C} t)$$
$$V_{IN} = A_{IN} \sin(\omega_{IN} t)$$

$$i_{EAM,Sing} = \left( \frac{dC_{S1}}{dx} \ddot{x}V_{C} + C_{S} \dot{V}_{C} \right) + \left( C_{Pa} \dot{V}_{IN a} \right)$$

$$i_{EAM,Sing} = \left( (\alpha_{1} \cos(\omega_{IN} t) + \alpha_{2} \cos(2\omega_{IN} t))\alpha_{3} \sin(\omega_{C} t) + \alpha_{4} \cos(\omega_{C} t) \right) + \left( \alpha_{5} \cos(\omega_{IN} t) \right)$$

$$i_{EAM,Sing} = \left( \frac{1}{2} \alpha_{1} \alpha_{3} (\sin(\omega_{C} t + \omega_{IN} t) + \sin(\omega_{C} t - \omega_{IN} t)) \right)$$

$$+ \frac{1}{2} \alpha_{2} \alpha_{3} (\sin(\omega_{C} t + 2\omega_{IN} t) + \sin(\omega_{C} t - 2\omega_{IN} t))$$

$$+ \alpha_{4} \cos(\omega_{C} t) \right) + \left( \alpha_{5} \cos(\omega_{IN} t) \right)$$

(A.21)

The second consideration is that this biasing scheme can be applied to either the single-sided or differential electrode schemes of Figures A.12 or A.13. The only difference
is the inclusion of the \(2\omega\) forcing frequency. For easy application to either system, Equation A.21 is generalized using coefficients \(\alpha_{\#}\). Replacing these coefficients with terms using Equations A.19 or A.20 allows application to either biasing scheme.

Through trigonometric identities, it can be shown that the multiplication of two sinusoidal signals results in the formation of two sidebands of the original high-frequency signal, separated from the original signal by the frequency of the second. This can be seen in Equation A.21, with the additional feed-through of the carrier at \(\omega_C\) and the drive voltage at \(\omega_{IN}\). After converting to voltage, the spectrum of this signal in the frequency domain can be seen in the first line of Figure A.14 (Output Voltage). Feed-through signals are depicted in red, \(\omega\) motional current in blue, and \(2\omega\) motional current in green.

Through the use of front-end electronics, which will be discussed in the following section, the output current of Equation A.21 is converted to voltage, represented by a gain \(K\) in Equation A.22. To help reduce noise in the system, it can be advantageous to pass the signal through a bandpass filter at this stage. Though not required, if used, this filter should only include the motional sidebands of the structure, and consequently, the feed-through carrier signal to remove the feed-through diving voltage and any DC offset.

\[
V_{\text{EAM,Sing}} = K_i V_{\text{EAM,Sing}} \\
V_{\text{Demod,Sing}} = V_{\text{EAM,Sing}} V_C + \alpha_6 V_C \\
V_{\text{Demod,Sing}} = \left( \frac{1}{2} \alpha_1 \alpha_3 K_A C \sin(\omega_C t) \sin(\omega_C t + \omega_{IN} t) \right. \\
\quad + \frac{1}{2} \alpha_1 \alpha_3 K_A C \sin(\omega_C t) \sin(\omega_C t - \omega_{IN} t) \\
\quad + \frac{1}{2} \alpha_2 \alpha_3 K_A C \sin(\omega_C t) \sin(\omega_C t + 2\omega_{IN} t) \\
\quad + \frac{1}{2} \alpha_2 \alpha_3 K_A C \sin(\omega_C t) \sin(\omega_C t - 2\omega_{IN} t) \\
\quad + \alpha_4 K_A C \sin(\omega_C t) \cos(\omega_C t) \bigg) + \alpha_6 A_C \sin(\omega_C t) \\
\quad + \left( \alpha_3 K_A C \sin(\omega_C t) \cos(\omega_{IN} t) \right) \tag{A.22}
\]

\[
V_{\text{Demod,Sing}} = \left( \frac{1}{4} \alpha_1 \alpha_3 K_A C \cos(-\omega_{IN} t) - \cos(2\omega_C t + \omega_{IN} t) \right) \\
\quad + \frac{1}{4} \alpha_1 \alpha_3 K_A C \cos(\omega_{IN} t) - \cos(2\omega_C t - \omega_{IN} t) \\
\quad + \frac{1}{4} \alpha_2 \alpha_3 K_A C \cos(-2\omega_{IN} t) - \cos(2\omega_C t + 2\omega_{IN} t) \\
\quad + \frac{1}{4} \alpha_2 \alpha_3 K_A C \cos(2\omega_{IN} t) - \cos(2\omega_C t - 2\omega_{IN} t) \\
\quad + \frac{1}{2} \alpha_4 K_A C \sin(2\omega_C t) \bigg) + \alpha_6 A_C \sin(\omega_C t) \\
\quad + \left( \frac{1}{2} \alpha_5 K_A C \sin(\omega_C t + \omega_{IN} t) + \sin(\omega_C t - \omega_{IN} t) \right)
\]
The signal must then be demodulated by multiplying by the carrier frequency, \( V_C \), as shown in Equation A.22. In addition to this ideal response, during this demodulation it is assumed the carrier signal feeds-through to the output signal, as well, with coefficient of magnitude \( \alpha_6 \). The derivation of Equation A.22 shows how the motional signals split once again to higher frequency around \( 2\omega_C \), as well as their original frequencies at \( \omega_{IN} \). This effect is visually shown in the spectral density plots of Figure A.14 (Demodulation).

The final step is to ignore the higher frequency responses by passing the signal through a low-pass filter, as displayed in Figure A.14 (Low-Pass Filter).

Through the use of this method, parasitic feed-through signals can be effectively removed from the observed dynamic response. When implemented with the 2-port, differential electrode scheme, this can also include a clean, single-frequency actuation force.

### A.2.8 Front-End Electronics

As it has been shown in previous sections, MEMS resonators can be capacitively forced through a voltage input, with the resulting velocity of the proof mass detected with a current output. While voltage can be measured directly, current measurements are less intuitive, requiring the user to convert the current to voltage before a measurement can take place. The simplest way to do this is feed the current directly into a resistor of known value, \( R_f \), and measure the voltage drop across the resistor, \( V_O - V_I = IR_f \). In this way, the value of the resistor is simply a gain placed on the measured current. This technique is shown in Figure A.15 (left). This solution is not ideal for electrostatic MEMS for three reasons: 1) The output current is small, 2) Parasitic capacitance, \( C_p \), is high, and 3) The power of the current source can be small. Because of the first issue, the value of the inserted resistor should be high to maximize the gain of the conversion. However, as this resistance increases, so does the time constant of the response, which is equal to: \( \tau = R_f C_p \). A high resistor gain combined with a high parasitic capacitance results in a measurement with a slow response time, which can be troublesome considering high resonator values or carrier frequencies \[158\]. Furthermore, the power required to convert
the current to voltage is reflected by: \( P = I^2R_f \). This power must be expended by the
current source, which may already be a very low power system, and therefore have a high
influence on the dynamics of the system.

An alternative approach is to use an operational amplifier to power this conversion
of energy, as shown in Figure A.15 (right). An operational amplifier is a type of inte-
grated circuit which consists a large number of transistors, along with a few resistors
and capacitors. The circuit contains three stages: 1) A differential amplifier with a
high-input-impedance, 2) A high-gain voltage amplifier, and 3) A low-impedance output
amplifier. A simplified schematic of this is shown in Figure A.16.

Using the variables identified in Figure A.16, the voltage source within the amplifier,
\( V \), is designed to always reflect the relation of Equation A.23.

\[
V = G_{OL}(V_+ - V_-)
\]  
(A.23)

where \( G_{OL} \) is the open-loop gain of the amplifier.

The ideal operational amplifier behavior can then be said to follow the following rules
[159]:

1. Open loop gain is infinite (\( G_{OL} = \infty \)).

2. Input impedance is infinite (\( R_{IN} = \infty \)), while output impedance is zero (\( R_{OUT} = 0 \)).

3. The input terminals, \( V_+ \) and \( V_- \), draw no current.

For this idealized behavior, feeding a current source into the operational amplifier
configuration shown in Figure A.15 (right) results in the circuit actively amplifying and
converting the current into voltage. This configuration is also known as a transimpedance
amplifier. Instead of passively forcing the current through a resistor, which relies on the

Figure A.16: Simplified schematic of an operational amplifier.
current source to expend the energy of conversion, this method allows the operational amplifier to do so, actively, at the cost of additional power. This is accomplished because the output of the amplifier, $V_O$, is continually set to a level equal to $V = G_{OL}(V_+ - V)$, as seen in Equation A.23. When $G_{OL} = \infty$ and $V_+$ is grounded, this results in the amplifier outputting any current or voltage necessary to make $V_-$ equal to ground. Because resistor $R_f$ is shorting this output to the negative terminal, any positive output voltage will force current backwards, across $R_f$. Therefore, when there is zero output from the current source, $V_O$ will be zero because there is nowhere for this current to flow (Rule 3: The terminals draw no current). When the output of the current source is positive, the operational amplifier will actively decrease $V_O$ to draw the current over $R_f$ so that Rule 3 will remain true. Because the current is drawn over the resistor by the operational amplifier, and not forced to by the current source, the time constant is no longer a function of the parasitic capacitance of the MEMS, $C_p$; the feed-through resistance, $R_f$, is free to be raised to maximize the conversion gain; and the current source expends no additional energy to accomplish this, leaving the MEMS dynamics uninfluenced.

Quantitatively, the output of the ideal operational amplifier, $V_O$, can be expressed by Equation A.24.

$$V_O = -R_f I$$ (A.24)

Unfortunately, this idealized behavior is an oversimplification. Operational amplifiers have a finite gain that varies over frequency, with impact as shown in Equation A.25. When this is combined with a high parasitic capacitance, $C_p$, the operational amplifier can experience a self-oscillation that appears in the voltage output.
\[
G_{OL} = G_{\text{max}} \frac{\omega_0}{\omega_0 + s} \\
\frac{V_O}{V_1} = -G_{\text{max}} \frac{\omega_0}{\omega_0 + s} \quad (A.25)
\]

To avoid these oscillations in the voltage output, a feedback capacitor, \(C_f\), is typically included in parallel to the feedback resistor, as shown in Figure A.17. This effectively adds a low-pass filter to the output which, when chosen correctly, can quell this resonance.

To determine the correct value of \(C_f\), a transfer function for the transimpedance amplifier can be formed by summing the current at node \(V_1\), then solving for the transfer function \(\frac{V_O}{I}\), as shown in Equation A.26 [160].

\[
\sum i_1 = 0 = I - V_1 C_p s - (V_1 - V_O) \left( \frac{1}{R_f} + C_f s \right) \\
I = -\frac{1}{G_{OL}} V_O C_p s - V_O \left( \frac{1}{G_{OL}} + 1 \right) \left( \frac{1 + C_f R_f s}{R_f} \right) \\
\frac{V_O}{I} = -G_{OL} C_p R_f s + (1 + G_{OL})(1 + C_f R_f s) \\
\frac{V_O}{I} = -G_{\text{max}} \frac{\omega_0}{\omega_0 + s} \cdot \frac{R_f}{C_p R_f s + (1 + G_{\text{max}} \frac{\omega_0}{\omega_0 + s})(1 + C_f R_f s)} \\
\frac{V_O}{I} = - \frac{G_{\text{max}} \omega_0}{C_p + C_f} \quad (A.26)
\]

The transfer function of Equation A.26 has two poles, which can be solved using the quadratic formula.

\[
s_{1,2} = -\frac{1 + \omega_0 R_f [C_p + (1 + G_{\text{max}}) C_f]}{2 R_f (C_p + C_f)} \left( 1 \pm j \sqrt{\frac{4 \omega_0 R_f (1 + G_{\text{max}})(C_p + C_f)}{(1 + \omega_0 R_f [C_p + (1 + G_{\text{max}}) C_f])^2} - 1} \right) \quad (A.27)
\]

The purpose of the feedback capacitor, \(C_f\), is to prevent oscillations in the gain of the transimpedance amplifier. This can best be accomplished by choosing a value for \(C_f\) that will critically damp the system, which occurs when the two poles of the system are equated: \(s_1 = s_2\). Using Equation A.27, this can be solved analytically; however, numerically solving for the poles while varying \(C_f\) is also a possibility. By plotting the motion in the complex plane, the optimal value of \(C_f\) can eventually be reached.
Figure A.18: Poles $s_1$ and $s_2$ in the complex plane as feedback capacitor, $C_f$, increments from 0.1 to 10 pF. The system is critically damped when $s_1 = s_2$.

An example system is used to demonstrate this effect in Figure A.18, where poles $s_1$ and $s_2$ are plotted in the complex plane for various values of $C_f$. For operational amplifier OP177, a feedback gain of 1 MΩ, and parasitic capacitance of 10 pF, the value of feedback capacitance that critically damps the system is 3.5 pF.

The impact of this optimal feedback capacitance can be observed when plotting the magnitude and phase of the transfer function of the system, Equation A.26, both with and without this feedback capacitance. This is shown in Figure A.19 for the example system. Without the feedback capacitor, a resonance can be seen at approximately 100 kHz; however, this peak completely disappears after the feedback capacitor is installed.

A single transimpedance amplifier is capable of both converting the current output of a MEMS sensor to voltage and amplifying it. For a single-sided channel, assuming the amplification is great enough, this is all that is required; the voltage can be read directly. For a differential channel, because there are two output signals coming from the MEMS sensor at the same frequency but separated by 180°, two transimpedance amplifiers are necessary; one for each signal. These two voltage outputs must then be recombined to form one output.

There are a number of ways to combine these two signals. Because the two signals are of the same frequency and 180° separated, each signal is simply the negative of its pair: $V_1 = -V_2$. This allows each signal to be fed through separate resistors and directly into opposing terminals of an operational amplifier, as shown in Figure A.20 (left). By placing a feedback gain around the negative input terminal, along with a grounding gain on the positive input terminal, the two signals are subtracted and amplified by $\frac{V_2}{V_2 - V_1} = \frac{R_2}{R_2}$.

The issue with this approach is that the input impedances to the terminals is both low.
and asymmetric: $Z_1 = R_2$ and $Z_2 = R_1 + R_2$ [161]. This results in signal loss and an asymmetric combination of input voltages $V_1$ and $V_2$, defeating the purpose of differential detection.

A second approach is to build on the structure of Figure A.20 (right) and buffer both inputs to the amplifier with additional operational amplifiers. This increases and equalizes the input impedance to the circuit, where both $Z_1$ and $Z_2$ are equal to the impedance of the operational amplifiers, resulting in a symmetric subtracting of the two differential input signals. Because each resistor comes with a pair, there is a possibility that the values of these resistors are not exactly equal. In this case, there will be asymmetric amplification which will be a function of the difference. This may be an issue when building such a circuit from discrete components; however, when all operational amplifiers and resistors are fabricated on the same integrated circuit, the resistance asymmetries can be minimal [161]. The challenge here is that the gain of can not be modified post-fabrication, resulting in specific chips being fabricated with fixed gains.

To solve this challenge, a third approach is shown in Figure A.21. By connecting the feedback loops of the buffering amplifiers with resistors, the gain of the circuit becomes: $\frac{V_2 - V_1}{V_2 - V_1} = (1 + \frac{2R_3}{R_G}) \frac{R_1}{R_2}$. Assuming symmetry of the other resistor pairs, gain can be manipulated by modifying a single resistor: $R_G$. This configuration represents the classic instrumentation amplifier, representing a single integrated circuit with high input impedances, which subtracts two input signals and amplifies it by a modifiable gain.

A single instrumentation amplifier can be an excellent method of combining the differential output from a MEMS resonator, after converting to voltage using transimpedance...
Figure A.20: Simplified instrumentation amplifiers displaying the design progression of the classic design, shown in Figure A.21. Single operational amplifier with low input impedance (left) and a design with two additional buffering amplifiers, but fixed gain (right).

Figure A.21: Basic schematic of the classic instrumentation amplifier using three separate operational amplifiers, three sets of symmetric resistors, and a programmable gain resistor.
amplifiers. The combined schematic for this detection scheme is shown in Figure A.22, where $V_1$ and $V_2$ are the differential current outputs from the MEMS resonator and $V_O$ is the voltage output after signal combination. The gain of this circuitry, $G_{sys}$, is a result from both the individual transimpedance gains, as well as the gain of the instrumentation amplifier during recombination.

Instrumentation amplifiers are not only useful for detection, but also signal conditioning when delivering actuation signals to the device. Depending on the AC signal generator used, the total voltage range is generally limited to ±5 to 10 Volts. When adding a DC offset to this value as well, if can even further limit the actuation force generated on the chip. For this reason, mixing AC and DC signals using analog components can be useful for generating larger actuation forces; an instrumentation amplifier can be used to amplify the AC component, while passive components can inject a DC offset.

Differential actuation also requires two AC signals with the same amplitude and frequency, but shifted in phase by 180°. This can easily be accomplished using the output of a single AC signal generator and two instrumentation amplifiers with identical amplification, as shown in Figure A.23.

The outputs of the instrumentation amplifiers produce AC signals with opposite gains. These signals are then injected with a DC offset using a capacitor and resistor in series with a DC voltage supply. Each resistor / capacitor set ($R_1$ and $C_1$) of Figure A.23 serves as a passive high-pass filter, with cut-off frequency of $f = \frac{1}{2\pi R_1 C_1}$. In practical terms, capacitor $C_1$ prevents a DC current flow, allowing only certain AC signals to pass. Voltage $V_{DC}$ injects a DC bias to nodes $V_1$ and $V_2$; however, resistor $R_1$ serves to resist this
flow of current, slowing the injection. The interaction between $R_1$ and $C_1$ produces the cut-off frequency. When the two channels are balanced, the DC power supply produces a steady stream of current; the fluctuations in the individual channels compensate one another.

A.2.9 Self-Resonance Electronics

In the previous section, front-end electronics were presented for both forcing the resonator and observing its velocity, both expressed in terms of a sinusoidal voltage. One of the prime advantages of vibratory MEMS devices that reduces C-SWaP is through the ability to condition the output signal and use it to drive the device. This is known as self-resonance, and is possible through the use of two additional components: a phase-delay circuit and a form of amplitude control. The feedback signal at various stages of this processing is shown in Figure A.24.

As was provided in Equation A.9, a forced harmonic oscillator which is under steady-state conditions (along with being driven at the natural resonance frequency with arguably high Q-factor) has a displacement that behaves in-phase to the forcing condition. The time derivative of this displacement produces velocity, which can be shown to have a relative 90° phase delay. Because the output current of the resonator is derived from velocity and not displacement, this phase delay must be compensated before attempting to implement the feedback loop. This can easily be accomplished using a phase-delay circuit with an offset of approximately 90°. Please note that phase delays of other electronic components, or within the resonator itself, can reduce this value and impart variability over time. Including a form of compensation for this drift may be necessary.

If this signal was fed directly into the AC drive electronics, it would be an unstable
Figure A.24: Input and output sinusoidal signals from a MEMS resonator, displaying the necessary signal processing for self-resonance.

system: the voltage would either be too low to oscillate the resonator, or too high and cause the resonator to reach its maximum amplitude and collide with itself. An equilibrium point does exist to allow oscillations at a fixed amplitude of motion; however, it is an unstable equilibrium point, where small deviations quickly move the system to one of two extremes. To stabilize this point, an additional feedback loop must be constructed. These feedback loops are known as automatic gain control (AGC).

Several types of AGCs exist, but they can generally be classified into two categories: coherent and noncoherent. Coherent AGCs rely on synchronization of the incoming signal to estimate amplitude, while noncoherent AGCs do not require this timing information. With this additional information, coherent AGCs generally have higher performance than noncoherent types, but require additional circuitry and power. Coherent AGCs will be discussed in more detail in the following section, so for now, a noncoherent type will be assumed.

A common approach for noncoherent AGCs utilizes an envelope detector for observing amplitude: a circuit consisting of a diode, capacitor, and resistor. A properly designed envelope detector follows the leading edge of a voltage input and disperses this energy at a slower rate than the input signal itself. In terms of a sinusoidal input, this allows instantaneous input voltages which are higher than the voltage of the envelope detector to feed energy to it; however, for instantaneous voltages less than the voltage of the envelope detector, energy leaks out at a constant rate. This effectively allows the detector to track the amplitude of the input signal without need for synchronization, but at the cost of bandwidth and noise. The output of the envelope detector is then used as the control of a variable gain amplifier (VGA), which is used to amplify the phase-delayed signal.
Figure A.25: Control scheme for self-resonance implementation using noncoherent AGC.

Implementation of this technique is shown in Figure A.25, with references to the signals shown in Figure A.24.

This control scheme functions by first feeding the signal to two parallel control paths: one for manipulating amplitude, and a second for manipulating phase. The amplitude control path passes the signal into an amplitude detector, which removes the resonator frequency to track only changes in amplitude, such as through the envelope detector previously described. This value is then subtracted from a constant voltage set point, $V_0$, and fed into a PID (proportional-integral-derivative) controller, which continually responds to the amplitude error signal, $e(t)$.

The phase control path first passes the signal through a phase-delay circuit to compensate for the $90^\circ$ phase shift between the drive and sense of the resonator. The resultant signal is then fed into a variable gain amplifier (VGA), which is controlled by the output of the amplitude PID controller. This adjusts the gain of the phase-delayed signal to the required value for constant amplitude resonance. This resultant signal is then fed back into the resonator and used as the AC component for driving the device.

Please note that this technique does not directly measure mechanical amplitude, but only the influence of the mechanical amplitude after passing through the read-out electronics. A direct measurement of mechanical amplitude has been shown to be possible through the use of the nonlinear effects of parallel plate electrodes [162]; however, this requires high amplitudes of motion and signal modulation, which adds additional circuitry to the device.

Instead of using a type of amplitude detection and control, it is also possible to simply limit the supply voltage of the read-out amplifiers to limit the mechanical displacement.
A.2.10 Phase and Amplitude Control

In the previous section, the schematic of a noncoherent AGC implementation was presented and discussed for the purpose of driving an electrostatic resonator into self-resonance. In this section, a coherent AGC approach is taken. In comparison, there are numerous benefits to this approach: 1) Increased robustness to frequency drifts, 2) Reduced phase noise and jitter, and 3) Reduced amplitude noise. In addition to this, there is also one major drawback to consider: additional circuitry which raises the C-SWaP of the device, which includes an additional oscillator of higher stability than the primary resonating structure. Implementation of this technique is shown in Figure A.26.

Like the previous section, this approach first separates the output signal of the resonator to two paths: one for manipulating amplitude, and a second for manipulating phase. To extract the phase information, the signal is demodulated by a second oscillator of preferably higher stability than the resonator being driven, with a 90° phase delay. When the two frequencies are in phase, the output of this demodulation is zero; however, deviations result in a phase error signal which is then passed into a PID controller. The output of this controller is then used to modify the frequency of a voltage controlled oscillator (VCO). The VCO output is then used for the prior modulation, allowing the two signals to continually stay synchronized. Typically, the VCO has a higher stability than the primary resonator.

To extract amplitude of the resonator, a similar demodulation step is performed, though without the phase delay. The output of this demodulation produces the amplitude of the resonator, which is similarly subtracted from a given set point and fed into a PID controller.
controller, as shown in the previous section. The output of the PID controller is the magnitude of the AC driving voltage, which is then modulated by the high stability oscillator with a 90° phase delay, to condition it for use as the AC component for driving the device.

The phase control shown here is commonly refereed to as a phase-locked loop (PLL). This control loop not only enhances the stability of the primary resonator through an additional PID phase controller, but also is used to completely reconstruct the driving signal. Assuming the VOC has a higher stability than the primary resonator, this allows for the reduction of phase noise through manipulation of the PID controller gains.

Through the use of a PLL and AGC, the two degrees of freedom of the output of the resonator are separated, passed through PID controllers, then recombined using a high stability oscillator before being used to drive the device. In the previous section, only one of these degrees of freedom were actively controlled, and the resonator stability completely dictated the stability of the drive signal.

A.2.11 Analog-to-Digital Converters

While most sensor designs can be implemented using only analog components, this approach has additional challenges in terms of noise and drift. Analog signals are not only sensitive to external stimuli, such as temperature changes and magnetic fields, but intrinsic noise sources as well, such as thermal fluctuations in each additional component. Digital signals are nearly immune to these challenges, but at the cost of additional power and cost. The use of analog signal processing may be a perfectly acceptable process for low-performance devices in interest of minimizing C-SWaP, but high-performance devices typically require an analog-to-digital conversion immediately following amplification. A highly stable clock signal may also be a desirable addition when processing digital signals, for both the implementation of certain control algorithms, as well as increasing the sampling rate for a higher accuracy conversion.

There are many types of analog-to-digital converters (ADCs) with various advantages and disadvantages [163], the trade-offs typically involve: 1) The number of digital bits of data (effectively, the dynamic range of each data point), 2) The maximum speed at which the conversion can occur, 3) The power and size of the converter, and 4) The level of parasitic capacitance the converter introduces to the analog system. Because this work is focused on the application of such converters, only three different types will be discussed due to their relevance: 1) Flash ADCs, 2) Sigma-Delta ADCs, and 3) Voltage-to-Frequency ADCs.

Flash ADCs, also know as parallel or direct conversion ADCs, are the classic ADC architecture. They function by first establishing a voltage ladder, as shown in Figure A.27, with regards to some reference voltage, $V_{REF}$. This voltage determines the range of the conversion. The input voltage to be converted, $V_{IN}$, is then fed into a comparator at each established voltage level. The comparator outputs a binary signal, which is then fed into a digital encoder, converting the 'height' that the analog signal achieved on the voltage ladder to a binary output. The digital encoder is comprised of a series of binary logic gates to perform this conversion, the number of which is exponential versus the number of rungs in the voltage ladder.
Flash ADCs are the fastest type of ADC implementation, with the sampling rate only limited by the speed of the comparators and logic gates. However, this type does not scale well with respect to increasing the number of bits of data being encoded. The number of components in the digital encoder effectively doubles with each additional bit being encoded, which increases size and power consumption. The additional rungs in the voltage ladder also increase the parasitic capacitance on the analog signal. For these reasons, Flash ADCs are typically not designed past 8 bit architectures; however, are capable of sampling rates up to 1 GHz.

Sigma-Delta ADCs, also known as Delta-Sigma ADCs, were designed to increase the number of encoded bits at the cost of processing speed. Figure A.28 shows a schematic of a simple first-order Sigma-Delta ADC. This design assumes that the ADC is operating at a frequency much higher than the analog input voltage, $V_{IN}$, so that $V_{IN}$ can be assumed to be constant over a few clock cycles. The constant analog input is then fed through an integrator, the output of which is a ramp function. The slope of the ramp is a reflection of the magnitude of the analog input. Once a critical threshold is reached, the comparator outputs a pulse, which is fed both into a digital filter, as well as subtracted from the analog signal before the integrator. This instantaneously reduces the magnitude of the output of the integrator, but the value continues to ramp upwards due to the constant analog input. When this continues for a few cycles, the time between the digital pulses of the comparator becomes an accurate representation of the magnitude of the analog input. The number of these pulses is summed over a fixed interval of time, dictated by some external clock, and converted to binary. The namesake of this technique is due to
the pulses required to reset the integrator (Delta), as well as the summation of pulses before converting to binary (Sigma).

Figure A.28 shows a simple first-order Sigma-Delta ADC; however, there is considerable room for design modifications. Additional integrators can be added in series to the left of the system, with the comparator feeding to the input of each. The number of integrators indicates the order of the system and can reduce noise in the output. The comparator also does not have to be a binary impulse, but rather can have multiple levels (the example being a 2-level version). In addition, modifications can be made to the magnitude of the digital pulses as well as the length of time of the summation before converting to binary. These factors combined lead to a sliding scale of the number of bits to sampling rate, typically ranging from 14 to 24 bits and sampling rates up to 10 MHz [163], [164].

The final type of ADC that will be discussed are Voltage-to-Frequency ADCs. This type of ADC directly converts an analog voltage to a frequency, which is then read by a pulse counter and converted to a binary output [165]. A simplified schematic of this is shown in Figure A.29.

These types of ADCs typically can not compete with the performance of the other types, and many times are not even mentioned due to this fact; however, they do have some niche applications. For instance, frequencies are substantially more robust to noise than analog voltages [164]. When a remote, analog sensor is in a noisy environment and power requirements prevent digital conversion locally, a voltage-to-frequency converter can be placed remotely with the sensor. The signal can then be transmitted a long
distance with minimal interference, and converted to binary at a location without power constraints. The performance of Voltage-to-Frequency ADCs are typically limited by the voltage-to-frequency converter.

A.3 Influence of Noise on Inertial MEMS

Noise and drift in the output of a sensor are critical for determining the precision and accuracy of a measurement. The boundary between these two terms can become blurred at times, but generally speaking, noise is considered a stochastic process, while drift represents a steady change of a parameter with time, possibly due to some external stimulus. There have been great strides in identifying the fundamental behaviors of these sources of error, but there are still many factors that are not fully understood.

In this section, noise and drift will first be classified from a cross-disciplinary experimentalist point of view. The impact of these various classifications will then be assessed in reference to the output of inertial MEMS, followed by the identification and calculation of specific noise sources that have been identified as primary influences on MEMS sensors.

A.3.1 Noise and Drift Characteristics

Noise can arise from a variety of sources depending on the parameter being measured. Regardless of this, a useful method of classifying different types of noise exists across disciplines, which separates noise by its power spectrum. Each noise classification is identified with a unique color, the most common of which are white, pink, and red. Figure A.30 displays each of these types of noise, along with linear drift, with respect to both time (left) and frequency (right).

White noise is defined as having a uniform power spectral density and can be viewed
<table>
<thead>
<tr>
<th>Parameter</th>
<th>PSD, $S_y(f)$</th>
<th>Allan Dev, $\sigma(\tau)$</th>
<th>Phase Noise, $L(f)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear Drift:</td>
<td>$f^{-2}$</td>
<td>$\tau^{+1}$</td>
<td>$f^{-4}$</td>
</tr>
<tr>
<td>Red Noise (Random Walk):</td>
<td>$f^{-2}$</td>
<td>$\tau^{+1/2}$</td>
<td>$f^{-4}$</td>
</tr>
<tr>
<td>Pink Noise (Flicker):</td>
<td>$f^{-1}$</td>
<td>$\tau^{0}$</td>
<td>$f^{-3}$</td>
</tr>
<tr>
<td>White Noise:</td>
<td>$f^{0}$</td>
<td>$\tau^{-1/2}$</td>
<td>$f^{-2}$</td>
</tr>
<tr>
<td>Random Walk of Phase:</td>
<td>$f^{0}$</td>
<td>$\tau^{-1/2}$</td>
<td>$f^{-2}$</td>
</tr>
<tr>
<td>Flicker Phase Noise:</td>
<td>$f^{+1}$</td>
<td>$\tau^{-1}$</td>
<td>$f^{-1}$</td>
</tr>
<tr>
<td>White Phase Noise:</td>
<td>$f^{+2}$</td>
<td>$\tau^{-1}$</td>
<td>$f^{0}$</td>
</tr>
</tbody>
</table>

Table A.3: Power laws for various types of noise and drift with respect to power spectral density, Allan variance, and phase noise (only applicable to frequency measurements). The top group is applicable to any signal, while frequency measurements also include the bottom segment.

as a Gaussian probability spread around the specific measured value in the time domain. Pick noise, also known as flicker or $1/f$ noise, has a motion in the time domain that resembles the motion of a flickering flame; momentary bursts of motion away from the measurement, but does not contribute to long-term drift. Red noise can be formed by integrating white noise over time, where data point $(n)$ has a Gaussian probability distribution around $(n-1)$. For this reason, red noise is also considered ‘random walk’, because over time it can cause drift substantially away from the measurement. Finally, linear drift has a similar spectral density as red noise; however, consistently drifts in one direction.

Another method of examining noise is by taking an Allan deviation of the data. Allan deviation is the square root of the Allan variance, which can be calculated for discrete cluster times, $\tau$, using Equation A.28.

$$\sigma_y^2(\tau) = \langle \sigma_y^2(2, \tau, \tau) \rangle = \frac{1}{2} (\bar{y}_{n+1} - \bar{y}_n)^2$$

$$\sigma_y^2(\tau, M) = AVAR(\tau, M) = \frac{1}{2(M-1)} \sum_{i=1}^{M-1} (y_{i+1} - y_i)^2 \quad (A.28)$$

As it can be seen in Equation A.28, Allan deviation can be thought of as half of the standard deviation of the data, with mean changing every $\tau$ seconds. This is calculated by taking a data set of $M$ values, dividing by $\tau$ to form $N = \frac{M}{\tau}$ sequential bins of data, then calculating a standard deviation across the entire sample range with the use of the mean of each individual bin. Allan deviation can be used to: 1) Separate the noise in a signal into various classifications, and 2) Determine an optimal averaging time for minimum output variance. Table A.3 lists several types of noise, along with linear drift, and their associated power laws when observing their power spectral densities and presence in Allan deviation calculations.

An example of an Allan deviation is shown in Figure A.28 for a data set that contains identifiable levels of white, pick, and red noise, along with a linear drift. In addition,
Figure A.31: Allan deviation of a signal with multiple noise sources. In order from left to right: White noise ($\tau^{-1/2}$), pink noise ($\tau^0$), red noise ($\tau^{+1/2}$), and linear drift ($\tau^{+1}$) because this plot was generated as a linear combination of multiple noise elements, Allan deviations of the individual components are potted as well for the regions in which they dominate to illustrate their effects.

For one-dimensional signals, such as the current or voltage output of amplitude modulated devices, these three types of noise, plus drift, characterize the fluctuations in the signals well. When frequency is measured as the output of a device, not only is the measurement sensitive to the fluctuations of frequency (white, pink, red, and drift), but also to fluctuations in the phase of the signal (white, pink, and red). Note that a linear drift of phase is identical to that of a shift in frequency. Equation A.29 compares the output of AM and FM devices, before and after noise elements are included in the response.

\[
x_{\text{AM}}(t) = x(t)
\]

\[
x_{\text{AM}}(t) = x(t) + \gamma(t)
\]

\[
x_{\text{FM}}(t) = A \cos(\omega t)\]

\[
x_{\text{FM}}(t) = A(1 + \alpha(t)) \cos((\omega + \beta(t))t) + \gamma(t)
\]

\[
x_{\text{FM}}(t) = A(1 + \alpha(t)) \cos(\omega t + \phi(t)) + \gamma(t)
\]

where $\alpha(t)$ is amplitude noise, $\beta(t)$ is frequency noise, $\phi(t)$ is phase noise, and $\gamma(t)$ is a non-periodic noise.
Figure A.32: Phase noise of an oscillator versus frequency offset from the carrier. From left to right: Random walk of frequency \((f^{-4})\), flicker of frequency \((f^{-3})\), white frequency noise and random walk of phase \((f^{-2})\), flicker of phase \((f^{-1})\), and white phase noise \((f^{0})\).

Frequency and phase fluctuations \((\beta(t)\) and \(\phi(t))\) are indistinguishable when attempting to measure an instantaneous frequency \((\dot{\beta}(t) = \dot{\phi}(t))\); however, most of their noise characteristics are not because the output is separated by a multiple of time, \(t\). This shifts the slopes of the noise spectra of phase, as shown in Table A.3. When this table is used to examine frequency outputs, the top segment applies to frequency, while the bottom segment applies to the phase of the signal. A few items to note: 1) White frequency noise is inseparable from the random walk of phase across all measurement techniques shown, and 2) White phase noise is inseparable from flicker phase noise when viewed with the Allan variance technique. If such a separation is desired, a modified Allan variance equation can be used.

The right column of Table A.3 also shows a third technique used to characterize the various types of noise for frequency measurements, along with the corresponding power laws for each component. This technique represents a single-sided phase noise measurement that can be implemented in multiple ways, such as using a phase-locked loop or observing the beats between two identical resonators. The result displays the phase noise of the oscillator versus frequency offset from the carrier. The observed slopes of the response can be used to measure the influence of each noise type. An example of this type of measurement is shown in Figure A.32.

The single-sided phase noise, \(L(f)\), is equivalent to twice the power spectral density of phase noise, \(S_\phi(f)\). This is because \(L(f)\) only includes noise in one direction from the carrier when measuring against the offset frequency. \(S_\phi(f)\) can then be converted to the power spectral density of frequency through Equation A.30 [166].
\[ L(f) = \frac{1}{2} S_\phi(f) \]
\[ S_y(f) = \frac{f^2}{f_0^2} S_\phi(f) \]  

where \( S_\phi(f) \) is the power spectral density of phase, \( S_y(f) \) is the power spectral density of fractional frequency fluctuations, and \( f_0 = 2\pi \omega \) is the carrier frequency.

### A.3.2 Precision and Accuracy of AM and FM Outputs

The uncertainty of the output of a sensor can be characterized by accuracy and precision. Accuracy is generally determined by some long-term sensor drift, while precision is largely determined by white noise. To increase precision, and if the time is available, additional samples can be taken and averaged for each measurement. Because white noise represents a Gaussian probability distribution around a given value, it can be effectively removed with infinite averaging. Pink noise, however, serves as a limit to this technique, where additional averaging of the signal results in no further increase in precision. Depending on the various noise levels, continued averaging may even result in reduced precision due to the random walk of red noise, or drift in the signal. The Allan deviation technique, shown in Figure A.31, is an excellent representation of precision versus averaging time.

Accuracy of a measurement; however, is more difficult to characterize. Assuming the measurement is linear, it takes the form shown in Equation A.31.

\[ y = (SF \times x) + B \]  

where \( y \) is the measured output, \( x \) is the inertial input, \( SF \) is the output sensitivity to changes in the inertial input, and \( B \) is an offset.

Accuracy is directly relatable to the stability of scale factor, \( SF \), and bias, \( B \). These stabilities are typically categorized into two separate values: in-run stability and run-to-run stability. Because scale factor is a rate of change, this value can also have a dependance on the value of the inertial input. This is referred to as 'scale factor nonlinearity'.

The in-run stability of bias can be observed through the use of the Allan deviation technique. The random walk of red noise represents a stochastic drift in bias, while the linear drift represents a deterministic dependance on some external stimulus (though the dependance on this stimulus may have its own errors). The run-to-run stability of bias must be measured by cycling the power of the system and measuring bias for each iteration. Assuming measurements can be made under zero inertial input, or compensated in such a way as to reflect a zero inertial input, accuracy of bias can be easily measured.

Scale factor stability poses a greater challenge. Measuring scale factor requires comparing the output of a device with that of a known inertial input. For accelerometers, this can be accomplished by a 2\( g \) tip-over test, with the accuracy of the input depending...
on the accuracy of the Earth’s gravity field. For gyroscopes, as will be later discussed, this is not as simple. Assuming an inertial input is available with stability below that of the sensor, in-run and run-to-run scale factor stabilities can be calculated in a similar way as described for bias. Though, because reorientation of the sensor is required, continual measurement of the sensor is not possible. Finally, scale factor nonlinearity can be measured across the range of the inertial input, assuming stability across this range is maintained.

Frequency measurements have the same restrictions as amplitude measurements, but with the added complexity of phase errors as well.

### A.3.3 Linear Drift

Noise and drift in the output of a sensor directly determines the precision and accuracy of a measurement. Understanding the various sources of noise and drift is critical for minimizing the influence of these factors during sensor design, as well as when designing in-run calibration techniques. For a majority of applications, the precision of uncompensated silicon inertial sensors is dominated by white noise, while the accuracy is dominated by thermal drift.

There are many sources of deterministic accuracy fluctuations in MEMS sensors, including both bias and scale factor. Various sensor parameters can be sensitive to aging, shock, acceleration, vibration, magnetism, and perhaps most critically, temperature. Due to the high sensitivity of silicon’s modulus of elasticity, some form of thermal compensation is nearly required for any substantial degree of performance. There are many techniques used to accomplish this, including: 1) Post-processing of data using on-chip thermometers, 2) Actively tuning resonance frequency using electrostatic springs, in response to on-chip thermometers, 3) Passively compensating the thermal coefficient of frequency by creating composite sensors, and 4) Using two devices oriented in opposite direction to create differential detection, robust to common-mode influences. While thermal drift has been the main focus of compensation techniques due to its high level of impact, the mechanisms behind many of the additional influences are must less understood, including aging and nonlinearities due to large displacements.

### A.3.4 Thermomechanical White Noise

As the main influence of measurement precision, the minimization of white noise is another priority for inertial sensors. Most applications of inertial sensors involve rapidly changing inertial inputs, which prevent long-term averaging and makes the reduction of white noise even more critical. For traditional, amplitude modulated inertial MEMS, white noise can be classified into three major categories: 1) Thermomechanical noise, 2) Electronics noise, and 3) Quantization noise. Thermomechanical noise and electronics noise are largely derivative of the same source: thermal fluctuations. The only difference being that the first applies to the resonant structure, while the second applies to the front-end electronics before digitization can occur. Quantization noise occurs from converting data between analog and digital domains, and is a function of both processing
bits and sampling rate. The impact of these noise sources for typical MEMS devices will now be assessed.

Thermomechanical noise is one of many manifestations of the fluctuation-dissipation theorem, which states that there is a quantifiable relationship between the generalized resistance and fluctuations of the generalized forces in linear dissipative systems [167]. This relationship is provided in Equation A.32.

\[
S_w(\omega) = 2k_BTR \quad , \quad -\infty < \omega < \infty \\
S_w(\omega) = 4k_BTR \quad , \quad \omega \geq 0
\]  

(A.32)

where \( S_w(\omega) \) is the power spectral density of the generalized force, \( R \) is the generalized resistance, \( k_B \) is Boltzman’s constant in joules per kelvin, and \( T \) is temperature in kelvin.

This force is generated by random particle collisions due to thermal fluctuations. Given any group of particles in a state of non-zero thermal equilibrium, as a group, their average temperature remains constant. On the atomic scale, however, any given particle is capable of any thermal state, with probability as determined by the Boltzmann distribution. This random distribution of particle energy states are then capable of interacting with their surroundings to produce macroscopic effects. The generated force is constrained to having a white characteristic due to lack of external force, which prevents a directional bias from developing over time.

This effect was first experimentally observed in the form of electrical potential, eventually being coined as Johnson-Nyquist noise, or simply thermal noise [168], [169]. The discovered relation is consistent with Equation A.32, but with the additional conversion to an RMS noise voltage for observation. Because white noise power is technically infinite due to its uniform, infinitely bounded power spectrum, this conversion requires a finite sampling bandwidth. These relations are given in Equation A.33.

\[
S_\nu = \bar{\nu}_n^2 = 4k_BTR \\
\nu_n = \sqrt{\bar{\nu}_n^2 \Delta f} = \sqrt{4k_BTR\Delta f}
\]  

(A.33)

where \( S_\nu \) is the one-sided voltage power spectral density, or voltage variance per hertz of bandwidth, \( \nu_n \) is the RMS noise voltage, \( k_B \) is Boltzman’s constant in joules per kelvin, \( T \) is temperature in kelvin, \( R \) is resistance in Ohms, and \( \Delta f \) is the bandwidth in Hz over which the noise is measured.

However, the fluctuation-dissipation theorem does not only apply to electrical potential, but any linear system with resistance. This is further supported by the equipartition theorem. The equipartition theorem states that for any collection of energy storage domains in thermal equilibrium, each domain will have an average energy equal to \( \frac{1}{2}k_BT \). This theorem applies to energy stored in any domain: kinetic (\( \frac{1}{2}mv^2 \)), spring potential (\( \frac{1}{2}kx^2 \)), electrical potential (\( \frac{1}{2}CV^2 \)), rotational (\( \frac{1}{2}I\omega^2 \)), etc [170]. This theorem can be generalized using the Hamiltonian of the given coordinate system [171], which is shown in Equation A.34, along with a few more specific applications.
\[
\langle x_m \frac{\partial H}{\partial x_n} \rangle = \delta_{mn} k_B T
\]

\[
\frac{1}{2} k_B T = \langle \frac{1}{2} m v^2 \rangle = \langle \frac{1}{2} k x^2 \rangle = \langle \frac{1}{2} C V^2 \rangle = \langle \frac{1}{2} I \omega^2 \rangle = \ldots
\]

where \( H \) is the Hamiltonian energy function with \( x_m \) degrees of freedom, \( \delta_{mn} \) is the Kronecker delta, and \( \langle \cdot \rangle \) denotes an ensemble average.

The fluctuation-dissipation theorem and equipartition theorem serve as the fundamental theoretical foundation of thermomechanical noise. Any dissipative linear system in non-zero thermal equilibrium can be analyzed for thermal forcing effects by simply adding a force generator parallel to each source of damping with magnitude reflective of Equation A.32. This is a fundamental law of nature and can only be minimized through the reduction of temperature or the magnitude of the resistive element. Just as Johnson-Nyquist noise specifically refers to the thermal noise of electrical potential, thermomechanical noise specifically refers to the thermal noise of mechanical systems. The application to specific types of inertial sensors will be discussed later in their corresponding sections.

### A.3.5 Quantization White Noise

Quantization noise occurs when converting information between analog and digital domains. Many times this is ignored with the assumption that the sampling rate and number of digital bits are both large enough to perfectly capture the signal. For prototype analog MEMS sensors, where information is digitized only for control algorithm implementation or read-out purposes, this can generally be the case. However, when these devices transition to become commercial products, these factors must be minimized to reduce cost. For digital MEMS sensors, where the control and output of the device is reflected in a one bit control signal, quantization noise can become much more prominent.

When an analog input is digitized, quantization error is defined as the information that is not captured in the transfer, and is related to the least significant bit (LSB) of the digitization, also called the quantum (Q). After conversion, each sample has an error range of \( \pm \frac{1}{2} Q \). Assuming the error is uniform throughout this range and uncorrelated to the sampling rate, it can then be converted to an RMS quantization error, \( V_Q \), the square of which is the power of this error signal, shown in Equation A.35 [174].

\[
P_Q = V_Q^2 = \frac{1}{Q} \int_{-Q/2}^{+Q/2} x^2 \, dx = \frac{1}{Q} \left[ \frac{x^3}{3} \right]_{-Q/2}^{+Q/2} = \frac{Q^2}{2^3} + \frac{Q^2}{2^3} = \frac{Q^2}{12}
\]

\[
P_Q = \frac{Q^2}{12}
\]

\[
V_Q = \frac{Q}{\sqrt{12}}
\]
To find the signal-to-noise ratio, the input signal is assumed to be a full-scale sinusoid, or a sinusoid which occupies the full range of digitation. The number of quantums throughout the full range of such a sinusoid is equal to $2^b$, where $b$ is the number of bits in the digitization, and the full-range voltage equals $Q2^b$. Thus, the digitized sinusoid can be expressed as $V_x = \frac{1}{2}Q2^b \sin(2\pi f t)$, with RMS value of $\frac{Q2^b}{2\sqrt{2}}$.

\begin{align*}
DR &= 20 \log_{10} \frac{Q2^b}{Q} = 20 \log_{10} 2^b \approx 6.02 \cdot b \\
SNR_{\text{Uniform}} &= DR \\
SNR_{\text{Sine}} &= 20 \log_{10} \frac{Q2^b / 2\sqrt{2}}{Q / \sqrt{12}} \\
SNR_{\text{Sine}} &= 20 \log_{10} 2^b + 20 \log_{10} \sqrt{\frac{3}{2}} \approx 6.02 \cdot b + 1.76
\end{align*}

Equation A.36

Note that the above relation is only valid over the Nyquist bandwidth from dc to $f_s/2$, where $f_s$ is the sample rate. When the bandwidth of interest remains constant and sampling rate is increased, the quantization noise spreads out over a larger frequency range, effectively reducing the noise within the bandwidth of interest. This reduction in noise can be added as a gain to the SNR, as shown in Equation A.37

\begin{align*}
SNR_{\text{New}} &= SNR_{\text{Old}} + 10 \log_{10} \frac{f_s}{2 \cdot \text{BW}} \\
SNR_{\text{Sine}} &= 20 \log_{10} 2^b + 20 \log_{10} \sqrt{\frac{3}{2}} + 10 \log_{10} \frac{f_s}{2 \cdot \text{BW}}
\end{align*}

Equations A.36 and A.37 are based off of the assumptions of Equation A.35: Quantization noise is uniform and uncorrelated to sampling rate. Uniformity is generally a valid assumption; however, quantization noise can become correlated to the sampling rate for certain frequencies being examined. When sampling rate can be reflected as an integer multiple of the frequency being examined, quantization noise concentrates as harmonics of the detected signal, effectively shaping the noise into false responses [175]. However, this can easily be avoided by slight modifications to the sampling rate.

For analog MEMS sensors, an effective design strategy is to determine the level of noise resulting from the mechanical and electrical components, then simply design the digital conversion to contain quantization noise on the same order of magnitude, or slightly below. Such a strategy effectively minimizes cost and power of the sensor.

**A.3.6 Noise in Frequency Modulated Outputs**

As was previously shown, the time domain error in an oscillating signal can be considered as fluctuation in frequency or phase, depending on the conversion. For electrical
components, this error is also commonly known as jitter. One of the most well known expressions for describing the phase noise of an oscillator is Leeson’s equation [176].

All oscillators comprise of a resonator, driven with feedback from its output; this was previously discussed under the topics of self-resonance and phase / amplitude control. In order to more accurately model the phase error in an oscillator, Leeson assumed that the phase noise of the resonator is transferred, without attenuation, through the feedback loop, with frequencies deviations of up to half the bandwidth of the resonator. The half-bandwidth, as shown in Equation A.7, is given by $\frac{1}{2} \text{BW} = \frac{\omega}{2Q} = \frac{1}{\tau}$. This relation is justified simply due to the dynamic response of a damped harmonic oscillator, where the frequency resolution is limited by Q-factor.

While this realization does not offer any insight to the actual cause of the phase fluctuations, it does offer an explanation for how a given oscillator reacts to phase fluctuations that are in its feedback loop. This relation is also known as the Leeson effect, and can be seen in Equation A.38, with corner frequency $\frac{f_0}{2Q}$.

$$S_\phi(f) = \left[1 + \left(\frac{f_0}{2Qf_m}\right)^2\right] S_\psi(f)$$

(A.38)

where $S_\phi(f)$ and $S_\psi(f)$ are the power spectral densities of phase noise in the oscillator and feedback loop, respectfully.

It is important to note that Leeson’s initially published equation was an empirical model for the experimental phase noise that he was observing. As such, it also includes several other noise sources that were dominant in his experiments, but it does not constitute a complete theory. Using the terms previously defined in Table A.3, in addition to the Leeson effect (which contributes to white frequency noise), Leeson also included thermal fluctuations in the form of white phase noise, as well as including an empirical value, $f_c$, which is dictated by the flicker of frequency. Leeson’s equation is supplied as Equation A.39.

$$L(f) = 10 \log \left[\frac{1}{2} \left(\left(\frac{f_0}{2Qf_m}\right)^2 + 1\right) \left(\frac{f_c}{f_m} + 1\right) \left(\frac{Fk_B T}{P_s}\right)\right]$$

(A.39)

where $f_0$ is the output frequency, $Q_l$ is the loaded Q-factor, $f_m$ is the frequency offset, $f_c$ is the $1/f$ corner frequency, $F$ is the noise factor of the amplifier, $k_B$ is Boltzmann’s constant, $T$ is the temperature in kelvin, and $P_s$ is the oscillator output power.

Many of the fundamental sources of these time domain errors are still not fully understood, but a number have been identified. Considering the resonator itself, contributing factors include: thermal noise, phonon scattering by material defects, temperature fluctuations, vibration, stress, radiation, and molecule adsorption, among others. Because of these factors, environmental stability, along with purity of the resonator material and surface conditions are the primary focuses of high-stability oscillator design.

As mentioned, noise in the feedback circuitry can also influence this time domain jitter. Electrical jitter is typically classified into two categories: deterministic jitter, and random
jitter [177]. Deterministic jitter is defined as having a non-Gaussian probability density function, is always bounded in amplitude, and has specific causes. There are four general classifications of deterministic jitter: duty cycle distortion, data dependent, sinusoidal, and uncorrelated (to the data) bounded. A number of well known sources contribute to these classifications, including signal cross-talk, electromagnetic interference (EMI), and any other regularly occurring noise source.

Random jitter, however, is probabilistic and unbounded, allowing it to influence long term stability of the output. The primary contributing factors of random electrical jitter are believed to be thermal vibrations in the silicon transistors and passive components, resulting in random electron mobility, as well as imperfections in doping density or other process anomalies.

A number of signal processing techniques can be used to help mitigate jitter, such as through the use of filtering or phase-locked loops, with effectiveness strongly dependant upon design.

A.4 Gyroscopes

To detect rotation, nearly all commercial MEMS gyroscopes utilize the Coriolis force to couple energy from one vibratory axis to a second, as a linear function of the rotational rate of the system. Such devices are called Coriolis Vibratory Gyroscopes (CVGs) and are rate measuring devices. In most situations, the attitude of the sensor is the true figure of merit, therefore the rate output must be integrated digitally over time to produce the current angle. Because this integration can produce errors, a second type of gyroscope exists which measure angle directly. These devices are called Whole Angle Gyroscopes. Both types of devices share similar dynamics, with some devices being capable of serving both functions depending on the implemented control. In this section, the dynamics, control, and noise metrics of several different types of MEMS vibratory gyroscopes is presented and compared.

A.4.1 Newton’s Equations of Motion in a Non-Inertial Frame

To examine the effects of rotation on a point in space, first consider two independent frames of reference, shown in Figure A.33: inertial frame $O_{\xi\eta\zeta}$, and non-inertial frame $O_{xyz}$.

Four vectors are denoted in this space: three position vectors, and two rotation vectors. $P_0$ extends from the inertial frame $O_{\xi\eta\zeta}$ to point $\alpha$, $d$ extends from the inertial frame $O_{\xi\eta\zeta}$ to point $\alpha$, $P_1$ extends from the local frame $O_{xyz}$ to point $\alpha$, $\theta$ is the rotation of the local frame with respect to the inertial frame $O_{\xi\eta\zeta}$, and $\Omega$ is the time derivative of $\theta$. Each of these vectors exist in three-dimensional space.

A rotation matrix, $R$, is then used to convert the local vector, $P_1$, to the inertial reference frame, by use of $\theta$. By acknowledging that $P_0$ is equivalent to the summation of $d$ and $RP_1$, two time derivatives can be taken to determine the inertial acceleration experienced by point $\alpha$. 

Figure A.33: Point in a rotating, non-inertial reference frame.

\[ P_0 = d + R P_1 \]
\[ \dot{P}_0 = \dot{d} + R \dot{P}_1 + (\Omega \times R P_1) \]  \hspace{1cm} (A.40)
\[ \ddot{P}_0 = \ddot{d} + R \ddot{P}_1 + (\Omega \times R P_1) + 2(\Omega \times R \dot{P}_1) + (\Omega \times (\Omega \times R P_1)) \]

The components of \( \ddot{P}_0 \) reveal a number of terms: \( \ddot{d} \) is the acceleration of local frame \( O_{xyz} \), \( R \ddot{P}_1 \) is the acceleration of point \( \alpha \) with respect to local frame \( O_{xyz} \), \( \dot{\Omega} \times R P_1 \) is the Euler acceleration acting on point \( \alpha \), \( 2(\Omega \times R \dot{P}_1) \) is Coriolis acceleration acting on point \( \alpha \), and \( \Omega \times (\Omega \times R P_1) \) is Centrifugal acceleration acting on point \( \alpha \).

These terms can then be rearranged with respect to local frame \( O_{xyz} \), and converted to a new notation.

\[ R \ddot{P}_1 = \ddot{P}_0 - \ddot{d} - (\dot{\Omega} \times R P_1) - 2(\Omega \times R \dot{P}_1) - (\Omega \times (\Omega \times R P_1)) \]
\[ \ddot{X}_0 = A - (\dot{\Omega} \times X_0) - 2(\Omega \times \dot{X}_0) - (\Omega \times (\Omega \times X_0)) \]  \hspace{1cm} (A.41)

where \( \ddot{X}_0 \) is equivalent to \( R \ddot{P}_1 \), representing the influences of internal forces, and \( A \) is equivalent to \( \ddot{P}_0 - \ddot{d} \), representing the influences of external forces. Both terms are from the perception of local frame \( O_{xyz} \).

Because each of these terms are three-dimensional vectors, \( X_0 \) will be defined as \( X_0 = [x_0; y_0; z_0] \), along with similar subscript notation for vectors \( A \) and \( \Omega \). Expanding Equation A.41 reveals Equation A.42.

\[ \ddot{x}_0 = A_x + (\Omega_y^2 + \Omega_z^2)x + (\dot{\Omega}_z - \Omega_y \Omega_x)y + (2\Omega_z)\dot{y} - (\dot{\Omega}_y + \Omega_z \Omega_x)z - (2\Omega_y)\dot{z} \]
\[ \ddot{y}_0 = A_y - (\dot{\Omega}_z + \Omega_x \Omega_y)x - (2\Omega_z)\dot{x} + (\Omega_z^2 + \Omega_x^2)y + (\dot{\Omega}_x - \Omega_z \Omega_y)z + (2\Omega_x)\dot{z} \]  \hspace{1cm} (A.42)
\[ \ddot{z}_0 = A_z + (\dot{\Omega}_y - \Omega_x \Omega_z)x + (2\Omega_y)\dot{x} - (\dot{\Omega}_x + \Omega_y \Omega_z)y - (2\Omega_x)\dot{y} + (\Omega_x^2 + \Omega_y^2)z \]

314
Equation A.42 represents the full acceleration experienced by a point from a non-inertial frame of reference.

A.4.2 Ideal Transducer Dynamics for Slow Rotations \((\Omega \ll 2\omega_n)\)

In the previous section there had been no generalizations, making Equation A.42 a fully accurate representation of the acceleration acting on a point, with respect to a non-inertial frame of reference. In this section, Equation A.42 is simplified to derive the dynamics of a Coriolis Vibratory Gyroscope.

Coriolis Vibratory Gyroscopes consist of a two-dimensional, harmonic oscillator, as shown in Figure A.34. These two axes are represented by the \(x\)-axis and \(y\)-axis, each with their own independent parameters, such as resonance frequency.

The system shown in Figure A.34 is similar to that of the previous section, but with the addition of specifying the internal dynamics of the system, \(\ddot{x}_0\). Equation A.43 illustrates the simplified dynamics shown in Figure A.34.

\[
\begin{align*}
\ddot{x}_0 &= \ddot{x} + \frac{\omega_x}{Q_x} \dot{x} + \omega_x^2 x - \frac{F_x}{m} \\
\ddot{y}_0 &= \ddot{y} + \frac{\omega_y}{Q_y} \dot{y} + \omega_y^2 y - \frac{F_y}{m} \\
\ddot{z}_0 &= 0 
\end{align*}
\]  

(A.43)

where \(\omega_x = \left(\frac{k_x}{m}\right)^{1/2}\) and \(\omega_y = \left(\frac{k_y}{m}\right)^{1/2}\) are the natural frequencies, while \(Q_x = \frac{\omega_x m}{c_x}\) and \(Q_y = \frac{\omega_y m}{c_y}\) are the Q-factors along the \(x\)-axis and \(y\)-axis, respectively.

Before the combination of Equations A.43 and A.42, Equation A.42 can be further reduced. As can be seen in Equation A.42, the rotation of the reference frame couple
into the dynamics of the system in three forms: Euler acceleration ($\dot{\Omega}x$), Coriolis acceleration ($2\Omega \dot{x}$), and Centrifugal acceleration ($\Omega^2 x$). When comparing these terms, Coriolis acceleration becomes dominate when Equation A.44 is true.

$$\frac{1}{2} \Omega \ll \omega_n$$
$$\frac{1}{2} \dot{\Omega} \ll \omega_n$$  \hspace{1cm} \text{(A.44)}$$

where $\omega_n$ is the operational frequency of the transducer.

When the inequalities of Equation A.44 become equalities, the Coriolis acceleration is equal in magnitude to the Centrifugal acceleration and Euler acceleration, respectively. This obfuscates measurements because a single rate is no longer isolated in the dynamics, which can potentially increase cross-axis sensitivity or sensitivity to rotational acceleration.

For the purpose of this dissertation, any phenomena that influences the system, compared to competing terms, by three orders of magnitude or less ($< 0.1\%$) will be ignored. Therefore, for a gyroscope with resonance frequency of $\omega_n \geq 1 kHz$: the input rate along any axis must be less than 2 Hz ($\dot{\Omega} \leq 2 Hz$), and the input rotational acceleration is a function of rate ($\ddot{\Omega} \leq 2Hz \cdot \Omega$).

When including the assumptions of Equation A.44, as well as the internal dynamics of Equation A.43, Equation A.42 can be reduced to Equation A.45.

$$\ddot{x} + \frac{\omega_x}{Q_x} \dot{x} + \omega_x^2 x - \frac{F_x}{m} = A_x + (2\Omega_x) \dot{y} - (2\Omega_y) \dot{z}$$
$$\ddot{y} + \frac{\omega_y}{Q_y} \dot{y} + \omega_y^2 y - \frac{F_y}{m} = A_y - (2\Omega_z) \dot{x} + (2\Omega_x) \dot{z}$$
$$0 = A_z + (2\Omega_y) \dot{x} - (2\Omega_x) \dot{y}$$  \hspace{1cm} \text{(A.45)}$$

Equation A.45 can be further reduced with two additional assumptions: 1) that the $z$-axis of the resonator is constrained, and 2) acceleration is negligible. These two assumptions can be validated through design of the system.

By using bulk silicon micro-machining techniques, thick devices can be fabricated with a high aspect ratio. The suspensions of these devices consist of cantilever beams, where the stiffness is proportional to the thickness of the beam, cubed ($k \propto t^3$). In order to create a $z$-axis, out-of-plane stiffness that exceeds in-plane stiffness by three orders of magnitude, the suspension must be 10 times thicker than the in-plane width. This is easy obtainable by existing fabrication techniques, and requires an aspect ratio of $\geq 10$.

In the previous section, the acceleration of a single point in a non-inertial frame was derived. The same conclusion can be made if a second, or multiple points are added to the system. When considering two points with internal dynamics equal in magnitude and opposite in sign, a differential measurement can be made using both points. When these points experience a common external acceleration, they may both be influenced by
the acceleration; however, the effect on their relative motion remains unaffected. This is known as anti-phase motion and will be discussed later in greater detail. For now, this design strategy will be used as justification for neglecting the influence of acceleration on the system.

When applying these two assumptions to Equation A.45, the result is shown in Equation A.46.

\[
\ddot{x} + \frac{\omega_x}{Q_x} \dot{x} + \omega_x^2 x = \frac{F_x}{m} + (2\Omega_x) \dot{y} \\
\ddot{y} + \frac{\omega_y}{Q_y} \dot{y} + \omega_y^2 y = \frac{F_y}{m} - (2\Omega_z) \dot{x}
\]

Equation A.46 represents the complete, traditional dynamics of Coriolis Vibratory Gyroscopes, with the following assumptions:

1. Linear rigid body dynamics
2. The principal axes of mass, stiffness, and damping are aligned to local frame \( O_{xyz} \)
3. The principal axes of force is aligned to local frame \( O_{xyz} \)
4. The rigid body does not rotate
5. All mass is susceptible to both the \( x \)- and \( y \)-axis system dynamics
6. Out-of-plane, \( z \)-axis motion is negligible (Fabrication aspect ratio of \( \geq 10 \))
7. Acceleration is negligible (Anti-phase internal dynamics)
8. \( \frac{1}{2}\Omega \ll \omega_n \) (for \( \omega_n \geq 1kHz: \Omega \leq 2Hz \))
9. \( \frac{1}{2}\dot{\Omega} \ll \omega_n \) (for \( \omega_n \geq 1kHz: \dot{\Omega} \leq 2Hz \cdot \Omega \))

A.4.3 Stiffness and Damping Asymmetry

While for educational purposes, Equation A.46 provides a useful framework for the operation of Coriolis Vibratory Gyroscopes, it fails to take into account asymmetries between the two modes of resonance (\( x \)- and \( y \)-axis). For all but the most intricately fabricated gyroscopes, this is not a valid assumption. Some gyroscope designs even purposefully design the \( x \)- and \( y \)-axis with different natural frequencies for enhanced bandwidth; however, even symmetric designs are vulnerable to fabrication tolerances, which create some order of asymmetry. For this reason, assumption two (The principal axes of mass, stiffness, and damping are aligned to local frame \( O_{xyz} \)) from the previous section will now be questioned, and the equations of motion modified.

The gyroscope model is redefined in Figure A.35. In this model, assumption two from the previous section have been replaced with a less restricted assumption: a Rayleigh
damping model, which forces symmetry of the damping matrix. This is because Rayleigh damping defines the damping matrix as a linear combination of two symmetric matrices: the mass matrix and the stiffness matrix ($[C] = \alpha [M] + \beta [K]$). Both of these matrices are symmetric by physics: kinetic energy ($T = \frac{1}{2} \dot{q}^T [M] \dot{q}$) and strain energy ($V = \frac{1}{2} q^T [K] q$), respectively. Assuming rigid body dynamics, kinetic energy and strain energy will both be positive for any non-zero displacement $q$. These definitions are analogous to stating that both the mass and stiffness matrices are positive-definite, and thus must be symmetric.

All positive-definite matrices are symmetric, but not all symmetric matrices are positive-definite. For both types of matrices there exists an orthonormal basis which projects the matrix into a diagonal form. This orthogonal basis is being defined as the principal axes of the corresponding matrix and are always 90 degrees apart from one another. The diagonal matrix consists of the eigenvalues of the original matrix and represent the magnitudes along each principal axis. For the mass and stiffness matrices, these magnitudes will always be positive, because the matrix is positive-definite; however, the damping matrix is free to be zero or negative. Zero eigenvalues represents zero damping along the axis, while a negative value represents energy; both conditions are impossible in nature without adding energy to the system.

No assumptions are being placed on the stiffness and damping matrices, apart from Rayleigh damping, which allows their principal axes to be independently misaligned from the local frame, $O_{xyz}$, though $\theta_k$ and $\theta_c$, respectively. Furthermore, the magnitudes of the principal axes are unconstrained (i.e., $\|k_1\|$ is not required to equal $\|k_2\|$, and $\|c_1\|$ is not required to equal $\|c_2\|$). While constraining assumptions still exist for the mass matrix, forcing it to be represented by the identify matrix, $I$, multiplied by a scalar, $m$, it is included here for completeness.

Because the main interest is in the dynamics of the system, and not the discrete
material properties, the three misaligned axes of Figure A.35 can be reduced to two by instead considering the resonance frequency ($\omega_n = \sqrt{\frac{k_n}{m_n}}$) and decay constant ($\tau_n = \frac{2Q_n}{\omega_n} = \frac{2m_a}{c_n}$). These new parameters are shown in Figure A.36, while Equation A.47 places Equation A.46 into matrix form.

$$\begin{bmatrix} \ddot{x} \\ \ddot{y} \end{bmatrix} + \begin{bmatrix} \frac{2}{\tau_x} & 0 \\ 0 & \frac{2}{\tau_y} \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} + \begin{bmatrix} \omega_x^2 & 0 \\ 0 & \omega_y^2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1}{m} & 0 \\ 0 & \frac{1}{m} \end{bmatrix} \begin{bmatrix} F_x \\ F_y \end{bmatrix} + \begin{bmatrix} 0 & 2\Omega_z \\ -2\Omega_z & 0 \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} \quad (A.47)$$

Equation A.47 still makes the assumption that the $x$-axis and $y$-axis are aligned to the principal axes of frequency and decay. For this reason, Equation A.48 allows the principal axes of frequency, $P_\omega$, and decay, $P_\tau$, to be misaligned from the local frame, $O_{xyz}$, by $\theta_\omega$ and $\theta_\tau$, respectively, as shown in Figure A.36.
\[ R_{\theta_\omega} = \begin{bmatrix} \cos(-\theta_\omega) & -\sin(-\theta_\omega) \\ \sin(-\theta_\omega) & \cos(-\theta_\omega) \end{bmatrix} \]

\[ R_{\theta_r} = \begin{bmatrix} \cos(-\theta_r) & -\sin(-\theta_r) \\ \sin(-\theta_r) & \cos(-\theta_r) \end{bmatrix} \]

\[ P_\omega = \begin{bmatrix} \omega_1 & 0 \\ 0 & \omega_2 \end{bmatrix} \]

\[ P_\tau = \begin{bmatrix} \tau_1 & 0 \\ 0 & \tau_2 \end{bmatrix} \]

\[
\begin{bmatrix} \ddot{x} \\ \ddot{y} \end{bmatrix} + 2R_{\theta_r}^T P_\tau^{-1} R_{\theta_\omega} \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} + R_{\theta_\omega}^T P_\omega P_\omega^T R_{\theta_\omega} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1}{m} & 0 \\ 0 & \frac{1}{m} \end{bmatrix} \begin{bmatrix} F_x \\ F_y \end{bmatrix} + \begin{bmatrix} 0 & 2\Omega_z \\ -2\Omega_z & 0 \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix}
\]

(A.48)

The new terms of Equation A.48 are calculated and displayed in Equation A.49.

\[
2R_{\theta_r}^T P_\tau^{-1} R_{\theta_\omega} = \begin{bmatrix} \left(\frac{1}{\tau_1} + \frac{1}{\tau_2}\right) + \left(\frac{1}{\tau_1} - \frac{1}{\tau_2}\right) \cos(2\theta_\tau) & \left(\frac{1}{\tau_1} - \frac{1}{\tau_2}\right) \sin(2\theta_\tau) \\ \left(\frac{1}{\tau_1} - \frac{1}{\tau_2}\right) \sin(2\theta_\tau) & \left(\frac{1}{\tau_1} + \frac{1}{\tau_2}\right) - \left(\frac{1}{\tau_1} - \frac{1}{\tau_2}\right) \cos(2\theta_\tau) \end{bmatrix}
\]

\[
R_{\theta_\omega}^T P_\omega P_\omega^T R_{\theta_\omega} = \begin{bmatrix} \frac{\omega_1^2 + \omega_2^2}{2} + \frac{\omega_1^2 - \omega_2^2}{2} \cos(2\theta_\omega) & \frac{\omega_1^2 - \omega_2^2}{2} \sin(2\theta_\omega) \\ \frac{\omega_1^2 - \omega_2^2}{2} \sin(2\theta_\omega) & \frac{\omega_1^2 + \omega_2^2}{2} - \frac{\omega_1^2 - \omega_2^2}{2} \cos(2\theta_\omega) \end{bmatrix}
\]

(A.49)

By inserting the terms from Equation A.49 into Equation A.48 and expanding, the result is shown in Equation A.50.

\[
\ddot{x} + \left(\frac{1}{\tau_1} + \frac{1}{\tau_2}\right) + \left(\frac{1}{\tau_1} - \frac{1}{\tau_2}\right) \cos(2\theta_\tau) \ddot{x} + \left(\frac{\omega_1^2 + \omega_2^2}{2}\right) + \left(\frac{\omega_1^2 - \omega_2^2}{2}\right) \cos(2\theta_\omega) \ddot{x} = \frac{F_x}{m} + (2\Omega_z) \ddot{y} - \left(\frac{1}{\tau_1} - \frac{1}{\tau_2}\right) \sin(2\theta_\tau) \dot{y} - \left(\frac{\omega_1^2 - \omega_2^2}{2}\right) \sin(2\theta_\omega) \dot{y}
\]

\[
\ddot{y} + \left(\frac{1}{\tau_1} + \frac{1}{\tau_2}\right) - \left(\frac{1}{\tau_1} - \frac{1}{\tau_2}\right) \cos(2\theta_\tau) \ddot{y} + \left(\frac{\omega_1^2 + \omega_2^2}{2}\right) - \left(\frac{\omega_1^2 - \omega_2^2}{2}\right) \cos(2\theta_\omega) \ddot{y} = \frac{F_y}{m} - (2\Omega_z) \ddot{x} - \left(\frac{1}{\tau_1} - \frac{1}{\tau_2}\right) \sin(2\theta_\tau) \dot{x} - \left(\frac{\omega_1^2 - \omega_2^2}{2}\right) \sin(2\theta_\omega) \dot{x}
\]

(A.50)

To simplify Equation A.50, the following terms are introduced: \( \frac{1}{\tau} = \frac{1}{\tau_1} + \frac{1}{\tau_2}\), \( \Delta(\frac{1}{\tau}) = \frac{1}{\tau_1} - \frac{1}{\tau_2}\), \( \omega = \frac{\omega_1 + \omega_2}{2}\), \( \Delta \omega = \omega_1 - \omega_2\), \( \Delta \omega^2 = \frac{\omega_1^2 - \omega_2^2}{2}\), and \( \omega \Delta \omega = \frac{\omega_1^2 + \omega_2^2}{2}\). With the inclusion of these terms, Equation A.50 becomes Equation A.51.
\[ \ddot{x} + \left( \frac{2}{\tau} + \Delta \left( \frac{1}{\tau} \right) \right) \dot{x} + \left( \omega^2 + \omega \Delta \omega \cos(2\theta_{\omega}) \right) x \\
= \frac{F_x}{m} + (2\Omega_z) \dot{y} - \Delta \left( \frac{1}{\tau} \right) \sin(2\theta_{\tau}) \dot{y} - \omega \Delta \omega \sin(2\theta_{\omega}) y \]

\[ \ddot{y} + \left( \frac{2}{\tau} - \Delta \left( \frac{1}{\tau} \right) \right) \cos(2\theta_{\tau}) \dot{y} + \left( \omega^2 - \omega \Delta \omega \cos(2\theta_{\omega}) \right) y \\
= \frac{F_y}{m} - (2\Omega_z) \dot{x} - \Delta \left( \frac{1}{\tau} \right) \sin(2\theta_{\tau}) \dot{x} - \omega \Delta \omega \sin(2\theta_{\omega}) x \]

Equation A.51

A summary of the assumptions of Equation A.51 are shown below:

1. Linear rigid body dynamics
2. Rayleigh damping
3. The principal axes of mass are aligned to local frame \( O_{xyz} \) with consideration to external forces (i.e. \( F_x, F_y, \) and the Coriolis force).
4. The principal axes of force is aligned to local frame \( O_{xyz} \)
5. The rigid body does not rotate
6. All mass is susceptible to both the \( x \)- and \( y \)-axis system dynamics
7. Out-of-plane, \( z \)-axis motion is negligible (Fabrication aspect ratio of \( \geq 10 \))
8. Acceleration is negligible (Anti-phase internal dynamics)
9. \( \frac{1}{2} \Omega \ll \omega_n \) (for \( \omega_n \geq 1kHz: \Omega \leq 2Hz \))
10. \( \frac{1}{\Omega} \ll \omega_n \) (for \( \omega_n \geq 1kHz: \dot{\Omega} \leq 2Hz \cdot \Omega \))

**A.4.4 Rigid Body Rotation and Force Misalignment**

In the previous section, the effects of the misalignment of the mass, stiffness, and damping matrices on the dynamics of a Coriolis Vibratory Gyroscope were derived and discussed. In this section, the effect of force misalignment on the rigid body will be addressed, challenging assumptions four (The principal axes of force is aligned to local frame \( O_{xyz} \)) and five (The rigid body does not rotate) of the previous section.

Up to this point, only translational motion of the gyroscope has been considered and rotation has been neglected. One of the reasons for this is because rotation is highly design dependant, requiring information on the exact placement of springs for an accurate representation. It is also common to assume that the impact of rotation will be small compared to the designed, translational motion. In order to challenge this assessment, a
Figure A.37: Multi-axis resonator with stiffness and force not acting through the center of mass.

A simple gyroscope is presented in Figure A.37 with two translational axes in the $x$- and $y$-axes, but is also allowed to rotate in-plane around the $z$-axis with state $\theta$.

The resonator presented in Figure A.37 has three state variables: $x$, $y$, and $\theta$. Two springs are attached along each linear axis at the edge of the proof mass. When assuming that $\theta$ will be restricted to small angles, the equations of motion are given by Equation A.52.

\[
\begin{align*}
\ddot{x} + \omega_x^2 x &= \frac{L_2 \Delta k_x \theta}{2m} + \frac{F_x}{m} \\
\ddot{y} + \omega_y^2 y &= \frac{L_1 \Delta k_y \theta}{2m} + \frac{F_y}{m} \\
\ddot{\theta} + \omega_\theta^2 \theta &= \frac{L_2 \Delta k_x}{2I} x + \frac{L_1 \Delta k_y}{2I} y + \frac{F_\theta}{I}
\end{align*}
\] (A.52)

where $F_x = F_1 + F_2$ and $F_y = F_3 + F_4$ are the combined forces in the $x$- and $y$-axis; $F_\theta = (-F_1 + F_2) \frac{L_2}{2} + (-F_3 + F_4) \frac{L_1}{2}$ is the combined rotational force; $k_x = k_1 + k_2$ and $k_y = k_3 + k_4$ are the combined stiffnesses in the $x$- and $y$-axis; $\Delta k_x = k_1 - k_2$ and $\Delta k_y = k_3 - k_4$ are the stiffness asymmetries; and $\omega_x = \sqrt{\frac{k_x}{m}}$, $\omega_y = \sqrt{\frac{k_y}{m}}$, $\omega_\theta = \sqrt{\frac{L_2 k_x + L_1 k_y}{2I}}$ are the natural frequencies of the $x$- and $y$-axis, as well as $z$-axis torsion.

When $x$ is forced into a sinusoidal motion using $F_x$, and $F_y = F_\theta = 0$, energy is coupled between the $x$- and $y$-axis through the rotation of the proof mass. This rotation is induced by the stiffness mismatch between the multiple springs along each axis. When observing the $y$-axis, this rotation takes the form of two sinusoids at the actuation frequency, the form of which is shown in Equation A.53.
\[x = x_0 \sin(\omega_x t)\]
\[y = T_R \cos(\omega_x t) + T_Q \sin(\omega_x t)\] (A.53)

The values of \(T_R\) and \(T_Q\) are both non-zero, even when neglecting damping. They can be passively minimized by increasing the torsional resonance frequency, \(\omega_\theta\), or decreasing the stiffness mismatches along each translation axis, \(\Delta k_x\) and \(\Delta k_y\). Though neglected in this case, mismatches in damping along each axis can also play a role in these terms.

Active compensation is also possible, should there be more than one forcer acting through different forcing vectors along each axis. An example of this is shown in Figure A.37, where \(F_\theta\) can be chosen, such that \(F_\theta = -\frac{L_2 \Delta k_x}{2I} x - \frac{L_1 \Delta k_y}{2I} y\). This, of course, is assuming that each force is perfectly aligned to its corresponding axis.

When comparing the influence of rotation with respect to the overall dynamics of the system, it is small. This can be verified by further simplifying the dynamics of Figure A.37 by assuming \(L_1 = L_2 = L\), \(\omega_x = \omega_y = \omega\), \(\Delta k_y = 0\), and \(F_\theta = 0\). The equation of motion for \(\theta\), along with the solution can be seen in Equation A.54

\[
\ddot{\theta} + \omega_\theta^2 \theta = \frac{L \Delta k_x}{2I} x \quad (A.54)
\]

where \(\frac{1}{2} L \theta\) represents the maximum linear displacement due to torsion and \(x_0\) is the nominal displacement along the \(x\)-axis due to forcing.

The last term of Equation A.54 shows the inequality that must be true to validate that the maximum torsional displacement is less than 0.1% of the driven motion, allowing this state to be ignored in the system dynamics. Stiffness mismatching, \(\frac{\Delta k_x}{k_x}\), is typically on the order of 1-2%, which implies that the torsional resonance frequency must be approximately five times larger than the operational frequency, or higher, to validate this inequality.

Due to the highly design-dependant influence of torsion, this inequality should only be used as an approximation for choosing torsional resonance frequencies. Furthermore, even if this effect is small compared to the drive motion, it still has the potential to influence the final performance of the device by inducing motion along the sense-axis. This is noteworthy due to Equation A.53; because the influence of this effect is out-of-phase with the resonance frequency, it is of particular interest, as will be later discussed. For this reason, its effects will be simplified in the generalized gyroscope equations of motion as acceleration terms of the form presented in Equation A.55.

\[x : T_x(\Delta k, \Delta c)\]
\[y : T_y(\Delta k, \Delta c)\] (A.55)
While purposefully designed rotational forces have the potential to compensate this effect, misalignment of the forcing vector can also induce unwanted rotation in the proof mass, as well as coupling between the two translational axes. Any rotational force induced in this manner is to be included in the terms of Equation A.55; however, direct translational coupling has yet to be included.

Figure A.38 depicts a model of this phenomena, where $F_1$ and $F_2$ are misaligned to the $x$- and $y$-axis by angles $\alpha$ and $\beta$, respectively. Due to these misalignments, each force enacts on the resonant proof mass along both translational axes. When including this effect, the new forces acting along the $x$- and $y$-axis are depicted in Equation A.56.

\[
R_{\alpha,\beta} = \begin{bmatrix}
\cos(-\alpha) & -\sin(-\beta) \\
\sin(-\alpha) & \cos(-\beta)
\end{bmatrix}
\]

\[
F = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix}
\]

\[
M_{xyz}^{-1} = \begin{bmatrix}
\frac{1}{m} & 0 \\
0 & \frac{1}{m}
\end{bmatrix}
\]

\[
R_{\alpha,\beta}F = \begin{bmatrix} F_1 \cos(\alpha) + F_2 \sin(\beta) \\ F_2 \cos(\beta) - F_1 \sin(\alpha) \end{bmatrix}
\]

\[
\begin{bmatrix}
\frac{F_x}{m} \\ \frac{F_y}{m}
\end{bmatrix} = M_{xyz}^{-1}R_{\alpha,\beta}F = \begin{bmatrix} \frac{F_x}{m} \cos(\alpha) + \frac{F_y}{m} \sin(\beta) \\ \frac{F_x}{m} \cos(\beta) - \frac{F_y}{m} \sin(\alpha) \end{bmatrix}
\]
By taking the effects described in Equations A.55 and A.56, and including them into the existing gyroscope dynamics of Equation A.51, the result is shown in Equation A.57.

\[\ddot{x} + \left(\frac{2}{\tau} + \Delta \left(\frac{1}{\tau}\right) \cos(2\theta_\tau)\right) \dot{x} + \left(\omega^2 + \omega \Delta \omega \cos(2\theta_\omega)\right) x = \frac{F_1}{m} \cos(\alpha) + \frac{F_2}{m} \sin(\beta)\]

\[+ (2\Omega z) \ddot{y} - \Delta \left(\frac{1}{\tau}\right) \sin(2\theta_\tau) \dot{y} - \omega \Delta \omega \sin(2\theta_\omega) y + T_x(\Delta k, \Delta c)\]

\[\ddot{y} + \left(\frac{2}{\tau} - \Delta \left(\frac{1}{\tau}\right) \cos(2\theta_\tau)\right) \dot{y} + \left(\omega^2 - \omega \Delta \omega \cos(2\theta_\omega)\right) y = \frac{F_2}{m} \cos(\beta) - \frac{F_1}{m} \sin(\alpha)\]

\[- (2\Omega z) \ddot{x} - \Delta \left(\frac{1}{\tau}\right) \sin(2\theta_\tau) \dot{x} - \omega \Delta \omega \sin(2\theta_\omega) x + T_y(\Delta k, \Delta c)\]  

(A.57)

A summary of the assumptions of Equation A.57 are shown below:

1. Linear rigid body dynamics
2. Rayleigh damping
3. The principal axes of mass are aligned to local frame \(O_{xyz}\) with consideration to external forces (i.e. \(F_1, F_2,\) and the Coriolis force).
4. The rigid body does not rotate (\(\frac{\Delta k_\mu}{k_\mu} \leq 0.001,\) or \(\omega_b \geq 5\omega_n\) for \(\frac{\Delta k_\mu}{k_\mu} \approx 1.5\%\))
5. All mass is susceptible to both the \(x\)- and \(y\)-axis system dynamics
6. Out-of-plane, \(z\)-axis motion is negligible (Fabrication aspect ratio of \(\geq 10\))
7. Acceleration is negligible (Anti-phase internal dynamics)
8. \(\frac{1}{2} \Omega \ll \omega_n\) (for \(\omega_n \geq 1kHz: \; \Omega \leq 2Hz\))
9. \(\frac{1}{2} \dot{\Omega} \ll \omega_n\) (for \(\omega_n \geq 1kHz: \; \dot{\Omega} \leq 2Hz \cdot \Omega\))

### A.4.5 Mass Decoupling and Angle Gain

Continuing with the development of the generalized gyroscope equations of motion, assumptions three (The principal axes of mass are aligned to local frame \(O_{xyz}\) with consideration to external forces) and five (All mass is susceptible to both the \(x\)- and \(y\)-axis system dynamics) of the previous section will now be examined.

While mass asymmetry has been briefly discussed in the previous sections, assumptions remained which forced it into the form of a scalar, \(m,\) multiplied by the identify matrix, \(I.\) To free this constraint, a multi-mass rigid body system will now be considered, shown in Figure A.39, and replace the single-mass system previously discussed.

The system of Figure A.39 consists of five separate masses. The central mass, \(m_c,\) is free to move along the two principal axes of the system, while the surrounding masses, \(\frac{1}{2}m_a\) and \(\frac{1}{2}m_b,\) are restrained to only a single principal axis. Torsional resonance of each
Figure A.39: Multi-axis resonator in a local frame of reference with decoupled mass matrix. The principal axes of mass are defined orthogonal, but allowed to be misaligned to the local frame.

Individual mass is assumed to be sufficiently high to validate assumption four (The rigid body does not rotate) of the previous section.

Figure A.39 depicts the foundation of a common CVG design strategy. Using this model, the diagonal elements of the mass matrix can now exhibit different values, based on the values of \( m_a \) and \( m_b \), which can serve to amplify or reduce the Coriolis force. Furthermore, because the mass along the \( x \)- and \( y \)-axis are different, this creates principal axes of mass which can now be misaligned to the local coordinate frame \( O_{xyz} \). This misalignment is represented by angle \( \theta_m \).

To account mass asymmetry in the free dynamics of the structure, the effect of this on the forcing terms (applied force and Coriolis) was neglected.

\[
M_{P,c} = \begin{bmatrix} m_c & 0 \\ 0 & m_c \end{bmatrix}
\]

\[
M_{P,tot} = \begin{bmatrix} m_a + m_c & 0 \\ 0 & m_b + m_c \end{bmatrix}
\]

When examining the right hand side of Equation A.47, Equation A.59 can be derived.

\[
\begin{bmatrix}
\frac{1}{m} & 0 \\
0 & \frac{1}{m}
\end{bmatrix}
\begin{bmatrix}
F_x \\
F_y
\end{bmatrix}
+ \begin{bmatrix}
0 & 2\Omega_z \\
-2\Omega_z & 0
\end{bmatrix}
\begin{bmatrix}
\dot{x} \\
\dot{y}
\end{bmatrix}
= R_{\theta_m}^T M_{P,tot}^{-1} R_{\theta_m} R_{\alpha,\beta} F + R_{\theta_m}^T M_{P,tot}^{-1} R_{\theta_m} C R_{\theta_m}^T M_{P,c} R_{\theta_m} \dot{X} + M_{xyz,tot}^{-1} R_{\alpha,\beta} F + M_{xyz,tot}^{-1} C M_{xyz,c} \dot{X} + \ldots
\]

\[
(A.58)
\]

326
Using a methodology similar to the process shown in Equation A.49, Equation A.60 has been formed.

\[
R_{\theta_m} = \begin{bmatrix}
\cos(-\theta_m) & -\sin(-\theta_m) \\
\sin(-\theta_m) & \cos(-\theta_m)
\end{bmatrix}
\]

\[
C = \begin{bmatrix}
0 & 2\Omega_z \\
-2\Omega_z & 0
\end{bmatrix}
\]

\[
M_{P,c} = \begin{bmatrix}
m_c & 0 \\
0 & m_c
\end{bmatrix}
\]

\[
M_{xyz,c} = R_{\theta_m}^T M_{P,c} R_{\theta_m}
\]

\[
M_{xyz,c} = \begin{bmatrix}
m_c & 0 \\
0 & m_c
\end{bmatrix}
\]

\[
M_{P,tot} = \begin{bmatrix}
(m_a + m_c) & 0 \\
0 & (m_b + m_c)
\end{bmatrix}
\]

\[
M_{xyz,tot}^{-1} = R_{\theta_m}^T M_{P,tot}^{-1} R_{\theta_m}
\]

\[
M_{xyz,tot}^{-1} = \begin{bmatrix}
\frac{\cos^2(\theta_m)}{m_a + m_c} + \frac{\sin^2(\theta_m)}{m_b + m_c} & \frac{\cos(\theta_m) \sin(\theta_m)}{m_a + m_c} & -\frac{\cos(\theta_m) \sin(\theta_m)}{m_b + m_c} \\
\frac{\cos(\theta_m) \sin(\theta_m)}{m_a + m_c} & \frac{\cos^2(\theta_m)}{m_b + m_c} + \frac{\sin^2(\theta_m)}{m_a + m_c} & \frac{\cos(\theta_m) \sin(\theta_m)}{m_a + m_c} \\
\frac{\sin^2(\theta_m)}{m_b + m_c} & \frac{\cos(\theta_m) \sin(\theta_m)}{m_a + m_c} & \frac{\cos^2(\theta_m)}{m_a + m_c} + \frac{\sin^2(\theta_m)}{m_b + m_c}
\end{bmatrix}
\]

\[
M_{xyz,tot}^{-1} = \begin{bmatrix}
\frac{1}{m_m} + \frac{1}{m_n} \cos(2\theta_m) & \frac{1}{m_m} \sin(2\theta_m) \\
\frac{1}{m_m} \sin(2\theta_m) & \frac{1}{m_m} - \frac{1}{m_m} \cos(2\theta_m)
\end{bmatrix}
\]

where \( \frac{1}{m_m} = \frac{1}{2(m_a + m_c)} + \frac{1}{2(m_b + m_c)} \), \( \frac{1}{m_m} = \frac{1}{2(m_a + m_c)} - \frac{1}{2(m_b + m_c)} \).

When \( \theta_m = 0 \):

\[
M_{xyz,tot}^{-1} R_{\alpha,\beta} F = \begin{bmatrix}
\frac{F_1}{m_a + m_c} \cos(\alpha) + \frac{F_2}{m_a + m_c} \sin(\beta) \\
\frac{F_2}{m_b + m_c} \cos(\beta) - \frac{F_1}{m_b + m_c} \sin(\alpha)
\end{bmatrix}
\]

\[
M_{xyz,tot}^{-1} CM_{xyz,tot} \dot{X} = \begin{bmatrix}
0 & 2\Omega \frac{m_c}{m_a + m_c} \\
-2\Omega \frac{m_a}{m_b + m_c} & 0
\end{bmatrix} \begin{bmatrix}
\dot{x} \\
\dot{y}
\end{bmatrix}
\]

(A.61)
\[
\ddot{x} + \left(2 \frac{1}{\tau} \cos(2\theta_x)\right) \dot{x} + \left(\omega^2 + \omega \Delta \omega \cos(2\theta_x)\right) x = \frac{F_1 \cos(\alpha)}{m_a + m_c} + \frac{F_2 \sin(\beta)}{m_a + m_c} \\
+ \left(2\Omega \frac{m_c}{m_a + m_c}\right) \dot{y} - \Delta \left(\frac{1}{\tau}\right) \sin(2\theta_x) \dot{y} - \omega \Delta \omega \sin(2\theta_x) y + T_x (\Delta k, \Delta c) \\
\ddot{y} + \left(2 \frac{1}{\tau} - \Delta \left(\frac{1}{\tau}\right) \cos(2\theta_x)\right) \dot{y} + \left(\omega^2 - \omega \Delta \omega \cos(2\theta_x)\right) y = \frac{F_2 \cos(\beta)}{m_b + m_c} - \frac{F_1 \sin(\alpha)}{m_b + m_c} \\
- \left(2\Omega \frac{m_c}{m_b + m_c}\right) \dot{x} - \Delta \left(\frac{1}{\tau}\right) \sin(2\theta_x) \dot{x} - \omega \Delta \omega \sin(2\theta_x) x + T_y (\Delta k, \Delta c)
\]

A summary of the assumptions of Equation A.62 are shown below:

1. Linear rigid body dynamics
2. Rayleigh damping
3. The principal axes of mass are aligned to local frame \(O_{xyz}\) with consideration to external forces (i.e. \(F_1, F_2,\) and the Coriolis force).
4. The rigid body does not rotate (\(\frac{\Delta k_x}{k_x} \leq 0.001,\) or \(\omega_k \geq 5\omega_n\) for \(\frac{\Delta k_x}{k_x} \approx 1.5\%\))
5. The satellite masses (\(m_a\) and \(m_b\)) are locked to the motion of a single-axis of the primary proof mass (\(m_c\)).
6. Out-of-plane, z-axis motion is negligible (Fabrication aspect ratio of \(\geq 10\))
7. Acceleration is negligible (Anti-phase internal dynamics)
8. \(\frac{1}{2} \Omega \ll \omega_n\) (for \(\omega_n \geq 1kHz: \Omega \leq 2Hz\))
9. \(\frac{1}{2} \Omega \ll \omega_n\) (for \(\omega_n \geq 1kHz: \dot{\Omega} \leq 2Hz \cdot \Omega\))

328
Appendix B

Silicon Device Fabrication

B.1 Introduction

The fabrication of released silicon structures can be a complicated subject. While a majority of the initial knowledge was gained through the integrated circuit industry, the field of MEMS quickly diverged as circuits continued to advance towards higher planar resolution, while MEMS pursued the missing third dimension. For high performance sensors, small errors when defining the thickness of the resonator can have large consequences.

Fabrication is further confounded by packaging requirements. Many MEMS sensors require vacuum packaging for enhanced performance, minimal stress imparted from the packaging method, rigid attachment for reducing substrate losses, yet high robustness to vibration and shock. This has led to a variety of fabrication and packaging techniques, with most commercial processes kept confidential.

In this chapter, there will be a brief overview of current and historical fabrication methodologies, an in-depth view of the silicon-on-insulator process utilized in the MicroSystems Laboratory at the University of California, Irvine, and finally the progress on an alternative approach towards a bulk dissolved wafer process for elimination of DRIE 'footing' asymmetry.

B.1.1 Fabrication Approaches

There have been many approaches toward device fabrication, a majority of which utilizing silicon as the resonator material due to the beneficial properties discussed in Appendix A. Initially, there were two basic approaches: surface or bulk micromachining. Surface micromachining is an additive process that utilizes a single-crystal-silicon wafer as a simple substrate for depositing and etching a series of thin films [186]. The resonator material typically consisted of a deposited layer of polysilicon, the properties of which could vary greatly depending on the deposition parameters. This variability also included intrinsic stress due to the high temperatures under which the material was deposited. This material variability, along with the fact that only thin layers could be deposited (on the order of a few microns [187]) led to alternative solutions.

In comparison, bulk micromachining is a subtractive process that carves the device
dimensions out of a single-crystal-silicon wafer. Due to a lack of techniques for anisotropically etching silicon, initial approaches were limited to utilizing chemicals that selectively etched certain silicon planes, such as KOH [188]. When using silicon wafers with a $<100>$ crystal orientation, this typically limited etching to the formation of concave trapezoidal cross-sections which prevented the creation of complex silicon structures.

It was not until the advent of deep reactive ion etching (DRIE), largely due to the Bosch process [189], [190], [191], that allowed the design of thick, complex single-crystal-silicon structures. As will be later discussed, DRIE involves exposing a masked wafer to at least two alternating environments: etching and deposition. By tuning the parameters of these two steps, trenches with relatively straight sidewalls can be etched into silicon, independent of crystal orientation, and with aspect ratios in excess of 30 [192]. When incorporated with a release step to suspend certain regions of the silicon structure, complex resonators can be formed.

### B.1.2 Bulk Micromachining Methods

There are two major challenges of bulk micromachining fabrication methods: 1) High aspect-ratio sensor definition, primarily for the formation of narrow capacitive gaps, and 2) Release of the resonant structure after etching. Initially, reactive ion etching (RIE) was largely used for the anisotropic etching of silicon, both single-crystal and polysilicon. For surface micromachining, this is a completely acceptable process to etch thin deposited films; however, the aspect ratio of this etching process, or the etch depth divided by the trench width, is generally limited to less than 10 [193]. Thin resonators are prone to a number of complications ranging from cross-axis sensitivity to electrode levitation; however, for many applications this aspect ratio is acceptable.

Another, less popular method of sensor definition utilized the selective doping of a single-crystal-silicon substrate. By then bonding the wafer to a second wafer with etched cavities, the substrate of the first wafer could be dissolved in a chemical bath, leaving behind only the doped regions of silicon [194]. Similar to RIE, this process was limited in not only aspect ratio, but also maximum device thickness, as determined by the dopant implantation method (10 to 20 $\mu$m, as shown in [194]).

Deep reactive ion etching (DRIE) solved many of these issues and quickly became a standardized method of fabrication. Since its initial conception, there have been a number of improvements made to the process for enhancing aspect ratio or improving sidewall angle or roughness. In the past, oxygen has been added to the passivation cycles to reduce sidewall bowing, while still achieving aspect ratios of 30 [195]. Process parameters have also been ramped during the progression of the etch for optimization as a function of etch depth. This work was able to produce aspect ratios that ranged from 70 to 100 with gaps from 2 to 5 $\mu$m [196]. And finally, most recently, epitaxial silicon deposition has been used post-DRIE to narrow trenches after etching, producing aspect ratios in excess of 500 [197].

However, high aspect ratio etching alone does not produce a resonator; release of the structure is also required. A number of methods have been used to achieve this. One of the first published papers of the DRIE process accomplished this by pre-etching a silicon substrate and bonding a second silicon wafer to encapsulate the gaps. The bonded
wafer was then etched using DRIE, exposing the cavities during the etching process, thus releasing the structures during etching [198].

Additional methods were also created that did not require wafer bonding, but rather utilized only a single silicon substrate, along with selective etching and passivation to form resonant structures. Examples include SCREAM (single-crystal reactive etching and metallization) [199], SIMPLE (silicon micromachining by single step plasma etching) [200], the porous silicon method [201], and SBM (surface/bulk micromachining) [202].

A similar process was also developed, coined as HARPSS (high aspect-ratio combined poly and single-crystal silicon). Though this process still utilizes polysilicon as the resonant structure, it is worthy to note due to the high aspect ratio capacitive gaps that can be obtained by the process. The process consists of first etching a mold into a silicon wafer and conformally coating this mold with a sacrificial layer of silicon dioxide, which is then selectively etched. Through further processing, polysilicon is then deposited into the molds and eventually the silicon dioxide is removed, forming narrow capacitive gaps, as defined by the oxide thickness. Using this process, capacitive gaps with aspect ratios of 50 were achieved [203].

Instead of beginning the process with a single silicon wafer, alternative approaches were also developed that began with wafers that had already been processed to include an embedded layer of silicon dioxide for release. These wafers, commonly referred to as silicon-on-insulator (SOI) wafers, could be formed using one of three methods: 1) Growing epitaxial silicon on a silicon wafer coated with oxide [204], 2) Implanting and annealing oxygen into a traditional silicon wafer [205], or 3) Bonding two silicon wafers with a buried layer of oxide [206]. The first two methods, respectively, produced silicon and oxide layers that were of poor quality compared to the Czochralski process and thermally grown oxide. The only disadvantage of the wafer bonding method is that there is the potential for voids [207]; however, this is easily avoidable with commercial processing.

Over time, several fabrication techniques evolved using commercially bonded silicon-on-insulator wafers. The black silicon method utilizes the buried silicon dioxide layer as part of a conformal coating of oxide around the resonant structure. The silicon substrate is then isotropically etched to release [193]. Alternatively, the buried silicon dioxide can simply be isotropically etched, once exposed due to etching of the device layer [208]. Modifications to this process can also be used for wafer-level sealing of devices [209].

In this work, commercially bonded silicon-on-insulator wafers are utilized to release thick silicon device layers that have been defined using DRIE. A detailed process flow follows, along with potential challenges and recommendations.

### B.2 Silicon-on-Insulator Fabrication Process

One of the basic methods of producing bulk silicon sensors is through the use of silicon-on-insulator (SOI) fabrication. This process utilizes the best features of both single-crystal silicon and silicon dioxide to produce resonant structures with excellent adhesion. Single-crystal silicon is useful for its predictable material properties, high stiffness, and potential for electrical conductivity, while silicon dioxide serves as an excellent adhesion layer, is electrically insulative, and can be selectively, and isotropically, etched.
B.2.1 Process Overview

The simplified silicon-on-insulator (SOI) fabrication process which was employed in this work is outlined in Figure B.1. Step one is the acquisition of an SOI wafer with a surface layer of silicon dioxide on the topside, but not on the bottom. The purpose of this oxide is to serve as a hard mask for device etching, while the backside must be pure silicon for adhesion of metal.

The second step is the blanket deposition of gold, with an adhesion layer of chrome. Step three patterns and etches this gold and chrome using lithography to create a selective die attachment area. Step four repeats the lithography process on the topside of the wafer with a mask that defines the features of the sensor. The silicon dioxide hard mask is then etched and the photoresist mask stripped. Step five is a deep reactive ion etch (DRIE) to carve the features of the sensor. And finally, step six releases the structure and removes the silicon dioxide hard mask using hydrofluoric acid (HF).

The process begins with the purchase of a number of silicon-on-insulator (SOI) wafers. Ultrasil has proved to be a reliable supplier of these wafers with consistent quality and capability to fulfill custom orders. Table B.1 shows the wafer properties that are typically requested for this process. At the time of this writing, minimum batches of 20 wafers are required for custom orders, with wafers costs of approximately 200-250 USD per unit.

Typical in-house SOI fabrication does not involve the deposition of front-size metallization. If front-side metallization is desired, the 1.5 µm of surface silicon dioxide is not required and wafers can either be ordered without this layer (for reduced cost), or this layer can be stripped with hydrofluoric acid (HF) before processing.

Another important note is that these wafers typically arrive in a large wafer carrier...
### Table B.1: SOI wafer properties.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Tolerance</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Top Surface</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Material:</td>
<td>Thermally-grown oxide</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thickness:</td>
<td>1.5</td>
<td>±0.05</td>
<td>µm</td>
</tr>
<tr>
<td>Finish:</td>
<td>Polished</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Device Layer</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diameter:</td>
<td>100</td>
<td>±0.1</td>
<td>mm</td>
</tr>
<tr>
<td>Orientation:</td>
<td>&lt;1-0-0&gt;</td>
<td>±0.5</td>
<td>deg</td>
</tr>
<tr>
<td>Flats:</td>
<td>Semi Std.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thickness:</td>
<td>100</td>
<td>±1</td>
<td>µm</td>
</tr>
<tr>
<td>Flatness:</td>
<td>&lt;2</td>
<td></td>
<td>µm</td>
</tr>
<tr>
<td>Type/Dopant:</td>
<td>P/Boron</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resistivity:</td>
<td>0.001 - 0.003</td>
<td></td>
<td>Ω − cm</td>
</tr>
<tr>
<td><strong>Buried Thermal Oxide</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thickness:</td>
<td>5</td>
<td>±0.25</td>
<td>µm</td>
</tr>
<tr>
<td><strong>Handle Wafer</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orientation:</td>
<td>&lt;1-0-0&gt;</td>
<td>±0.5</td>
<td>deg</td>
</tr>
<tr>
<td>Thickness:</td>
<td>500</td>
<td>±10</td>
<td>µm</td>
</tr>
<tr>
<td>Type/Dopant:</td>
<td>P/Boron</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resistivity:</td>
<td>1-10</td>
<td></td>
<td>Ω − cm</td>
</tr>
</tbody>
</table>

Table B.1: SOI wafer properties.
with all wafers exposed to each other. Wafers that arrive from Ultrasil are clean, therefore
do not require in-house cleaning before beginning the process. However, to ensure that
they remain clean before processing begins, the wafer container must be stored in a clean
environment, such as the INRF or MicroSystems cleanroom, and only opened in INRF
fume hoods. Only open this container once and load the wafers needed for processing
into individual wafer carriers. Individual wafer carriers are useful to not only track
the processing of each wafer, but also to reduce cross contamination. Once loaded into
individual holders, a marker can be used to write the batch and wafer number on the
outside of each holder for tracking purposes.

B.2.2 Backside Metal Deposition

The first step of the process is to pattern the backside of the wafer with chrome and gold
metallization. There are four reasons for this: 1) For a rigid, eutectic die attachment once
fabrication is complete, 2) To reduce the influence of die attachment stress, 3) To increase
the coefficient of friction on the backside of the wafers, and 4) To make a majority of
each silicon die transparent to infrared (IR) light.

A rigid die attachment is useful for reducing the degrees of freedom between the silicon
die and ceramic package, which can otherwise reduce the Q-factor of the device due to
anchor loss [210]. However, this rigid attachment can impart intrinsic stress during the
die attachment process which changes over temperature due to mismatches in the thermal
coefficient of expansion [7]. To benefit from the increased rigidity, while still minimizing
the influence of stress, a reduced die attachment area is utilized. In addition to this, it
has been observed that a 100% gold coating on the backside of silicon wafers has a lower
friction than the polished silicon itself. This can create issues when using certain pieces
of dry etching equipment due to wafers slipping during transfer (primarily, the FDRIE
at the UCLA Nanolab). By reducing the area of backside gold, the total coefficient
of friction is increased, thus reducing slippage. Finally, silicon and silicon dioxide are
transparent to certain wavelengths of light; however, gold is not. Using a camera that
detects this frequency of light, it is possible to inspect dies during their release to find
the proper release time on a die-by-die basis.

Wafer metallization can be completed in one of two ways: 1) Metallization before
patternning, or 2) Patternning before metallization. The first method coats the entire wafer
surface with metal, which is then patterned with a protective coating over the areas that
are meant to remain, and then soaked in a metal etchant. The second method patterns
the surface of the wafer with a special photoresist that develops more of the photoresist at
the wafer interface than at the interface with air. Metal is then deposited on this surface,
but due to the line-of-sight deposition of thermal evaporation, the undercut areas remain
free of metal. This creates broken metal traces where the patterning was place. The wafer
is then soaked in a photoresist stripper, which eventually dissolved all of the photoresist,
pulling away the unnecessary metal, and leaving only the desired metal traces.

While the second takes more time due to the stripping of the photoresist, it prevents
the wafer from unnecessary exposure to strong acids and can also result in improved
feature resolution. Because the features of the die attachment mask are large with respect
to standard lithography procedures and the only materials on the wafer are silicon and
silicon dioxide, the first method is being employed. The reduced feature resolution is even a benefit in case of pinpoint lithography errors.

The process begins by first stripping the native layer of oxide on the backside of the silicon wafers. When exposed to ambient conditions, silicon develops a thin layer of silicon dioxide on its surface. The thickness of this layer depends on the temperature, time, humidity, and oxygen content of the environment, but is generally below 1 nm [211]. To remove this layer, without significantly influencing the silicon dioxide hard mask on the front-side of the wafer, a 2% hydrofluoric acid (HF) solution is used, which has an etch rate of $\sim 5 \text{ nm/min}$ at room temperature ($21 - 25 \, ^\circ\text{C}$) [212].

Two baths should be made: one of 2% HF, and a second of DI water. The wafer should be placed in the HF bath with the backside place upwards, and allowed to soak for 15 seconds. Afterwards, the wafer is then rinsed with DI water in the sink and submerged in the DI water bath. When rinsing the wafer in the sink, observe how the water reacts to either side of the wafer. On one side, the frontside, there is still a thick layer of silicon dioxide, which is hydrophillic and causes the wafer to spread over the surface of the wafer easily and flow off the wafer. On the other side, the backside, the silicon dioxide has been stripped, resulting in the exposure of pure silicon, which is hydrophobic and causes the wafer to bead up into droplets before rolling off the wafer. This is a visual indicator that the native oxide was successfully stripped and is a useful indicator to track so that the correct side of the wafer is metalized. The purpose of the DI water bath is to dilute and remove any remanence of HF at the location the wafer is gripped with the tweezers. It also allows the user to rinse the tweezers in the sink.

After a few seconds, or once finished rinsing the tweezers, remove the wafer from the water bath and place the wafer inclined on a Texwipe. Grip the wafer at a low location and blow dry the wafer using a nitrogen air gun. To avoid contamination of the clean wafer by any particles on the tweezers, always blow dry from the wafer towards the tweezers. Also, only allow the Texwipe to come into contact with the wafer at one location on the edge. While the Texwipes are designed to produce low amount of particles, they still produce some particles and can contaminate the wafer if fully placed on them. Placing the wafer in contact with the Texwipe at one location is required to allow the water an avenue to travel from the wafer to the Texwipe. It also allows some increased stability when forcing the wafer the nitrogen air current. Caution must also be made when controlling the air flow; too much can break or cause the wafer to slip from the tweezers, while too little will not dry the wafer adequately. Once dry, return the wafer to the individual holder, record any issues, and continue with the next wafer and repeat.

Once all wafers have been stripped of native oxide, they are to be loaded into the thermal evaporator, along with the chrome and gold targets in their respective graphite crucibles. If the backside of the wafer is unknown, this can usually be identified by observing the edge of the wafer to determine which silicon layer is thicker. When suspending the silicon wafers in the evaporator, do not use Kapton tape. While many other users may use this, it can contaminate the wafers and be difficult to remove. Instead, there are a number of small metal clips that can be used to fix the wafers in place. Once loaded, close the machine and begin pumping the chamber to vacuum. Once the roughing pump
is complete and the high vacuum pump is opened, the chamber will take at least 1 hour to reach an acceptable level of vacuum. Generally, it is best to start this process late in the day, allow the vacuum to pump down overnight, then return in the morning to complete the deposition. This length of time does not affect the wafer, because native oxide can not reform without the presence of oxygen.

Once vacuum has been reached, proceed to deposit 50 nm of chrome, followed by 500 nm of gold. Gold is required for eutectic die attachment and chrome is necessary as an adhesion layer to allow the gold to stick to the silicon. In other words, gold does not stick to silicon; however, gold sticks to chrome, and chrome sticks to silicon. When completing the deposition, it is important to monitor the deposition rate. While there is some room for error, it is generally best to start and end the deposition of any material slowly, while only ramping up speed during the middle of the process. Deposition rate is directly related to uniformity across the wafer. By starting the deposition of any material slowly, a solid, uniform foundation is formed at the material interfaces which aids in adhesion. For this reason, the following deposition schedule is recommended, which optimizes deposition within the allotted time frame.

Chrome (Cr):

1. 0 to 10 nm, 0.01 nm/sec.
2. 10 to 40 nm, 0.03 nm/sec.
3. 40 to 50 nm, 0.01 nm/sec.

Gold (Au):

1. 0 to 10 nm, 0.01 nm/sec.
2. 10 to 50 nm, 0.05 nm/sec.
3. 50 to 450 nm, 0.1 nm/sec.
4. 450 to 500 nm, 0.05 nm/sec.

The total deposition time takes about 2.5 hours. To reduce this time, the rate of gold deposition can be increased during the bulk of the deposition (from 50 to 450 nm); however, it is not recommended to increase the deposition rate of chrome, considering the importance of its adhesion strength.

Once complete, the chamber will need at least 30 minuets to cool before the samples can be unloaded. The backside of the wafers should be fully metalized, as shown in Figure B.2.
B.2.3 Lithography on Metal

At this point in time, the front-sides of the wafers are significantly contaminated from loading them into the deposition machine, but the backside are not. There may be some particulates on the backside from loading and unloading into the evaporator, but the bulk of these can be removed with a simple solvent rinse. This is completed by holding the wafer with tweezers above a liquid container and flooding both surfaces with acetone. Before the wafer can dry, flood the wafer again with isopropyl alcohol (IPA), following by rinsing in the sink with DI water. When flooding the wafer, please note that the same rules apply with regards to contamination from the tweezers: always hold grip the wafer at the bottom and never flow liquid from the tweezers towards the center of the wafer. Once complete, blow dry the wafers with nitrogen, following the same procedure as outlined earlier.

The wafers should now be dehydrated to remove any moisture from the silicon and metal surface. This is performed by placing the wafers for 1 hour in an oven set to a temperature of 120 °C. Afterwards, the wafers are removed and briefly blow dried with nitrogen.

Immediately after dehydration, lithography is performed on the metallization. First, the wafer is placed on the chuck of the photoresist spinner. Spin the table by hand and observe the edge of the wafer to identify any eccentricity. By looking at one position fixed in space, you should see the edge of the wafer move laterally left and right as the table spins. The edge of the wafer must always be in the center of this oscillation. If the wafer is not centered, rotate the table until the edge of the wafer is at the maximum displacement away from the center of the chuck and use tweezers to gently push the edge of the wafer back towards the center to realign. Repeat until you can not visually identify any eccentricity. Engage vacuum and blow dry the wafer on the chuck. Open the bottle of AZ4620 photoresist and wipe the rim to ensure there are no dry particles. If many particles exist, do not use this bottle and purchase a new one from the INRF store. Pour the photoresist on the wafer, keeping in mind that too much is better than not enough. Wipe the rim of the bottle and quickly close. Spin the wafer using the following program:
1. 10 sec @ 500 rpm, 2 acc

2. 40 sec @ 5000 rpm, 6 acc

The above program should result in a $\sim 7 \mu m$ thick layer of resist. Soft bake for 20 min @ 90 $^\circ$C in the bake-out oven, and attempt to minimize the number of times the oven is opened to produce consistent results. Please note that you should have two sets of tweezers: one for dirty processes (or when photoresist is applied to the wafers), and one for clean. While this alone reduces wafer contamination, it is good practice to periodically clean both with solvents or RCA-1. Once the soft bake is complete, expose the wafer using the MA6 and die attachment mask with the following parameters:

1. Soft Contact: 25 $\mu m$.

2. Time: 36 sec.

A few notes: 1) Ensure that the light power during exposure is 10 $mW/cm^2$ (readout is on a box under the MA6), 2) As each wafer is exposing, keep the other wafers out of line-of-sight from the mask aligner to avoid additional polymerization, 3) Do not look at the UV light, and 4) When moving the wafers from the exposure room to the development room, keep wafers in the dark.

Another important thing to consider is that this is the first mask being applied to the silicon wafer; therefore, all subsequent masks should be aligned to this one. This means that this mask defines the orientation of sensors with respect to the silicon wafer. As shown in Appendix A, because $<1-0-0>$ wafers are being used, the angle of rotation of the wafer to the mask will play a role in the physical properties of the silicon. To insure uniformity, it is best to attempt to align the wafer flat to the mask as best as possible. The MA6 does not have this capability, therefore it must be done visually. When aligning the wafer to the mask, the MA6 has three controls: vertical, horizontal, and rotation. The rotation is the most difficult to adjust, because the misalignment between the wafer and mask is unobservable. After the first mask is applied this variable becomes observable, but the application of the first mask is blind. This can be done with reasonable accuracy by first moving the wafer up vertically until the wafer flat becomes visible through the lithography mask. Rotation can then be adjusted to attempt to line the flat to a straight horizontal line on the mask. The vertical position can then be moved back until the wafer just passes over the horizontal line on the mask. By moving the wafer vertically back and forth over this line, it becomes easier to see which end of the wafer flat disappeared first, and by how much. Continue to adjust the rotation until both ends of the wafer flat appear to disappear and reappear at once when oscillating the wafer vertically. Once the rotation is aligned, the horizontal and vertical alignment can be adjusted to place the features in the center of the wafer.

Once all wafers have been exposed, develop the photoresist by first preparing a bath of DI water and developer with a 3.5:1 ratio. Mix briefly with tweezers. Quickly drop the wafer into the mixture, with the photoresist facing upward. Do not agitate the solution.
Figure B.3: Backside of an SOI wafer after with patterned photoresist masking prior metallization. The pattern consists of a grid of assorted sizes.

during development, as this can influence development time and produce inconsistent results. Wait 2 minutes (only for backside mask). Quickly remove wafer and rinse with DI water in the sink. Briefly dry with nitrogen to remove a bulk of the water on the patterned side of the wafer and observe under a microscope. Check in 5 locations (center, top, bottom, left, right) and ensure that all of the photoresist in the trenches is removed. If photoresist remains, return wafer to the developer for 15 seconds and repeat as necessary. Please note that repeated emersion of the wafer leads to asymmetric development, so finding the optimal development time as wafers are processed is ideal. If features begin to debond, the wafer was either overdeveloped or exposure time / power was incorrect. Once complete, fully dry with nitrogen, as previously described. The resulting lithography on top of the metallization will appear similar to Figure B.3, though the mask features may be different.

B.2.4 Backside Metal Etching

With the metallization properly masked, the exposed metallization can be etched. Prepare three baths: one bath of gold etchant, one bath of chrome etchant, and one bath of DI water. Please note that the gold and chrome baths may already be prepared. The etchant solutions can be stored and reused many times, and are generally only disposed when they become contaminated or the etch rate has been reduced to unacceptable levels. Begin by quickly submerging the wafer into the gold etchant, with the features facing upward. Leave the wafer in the liquid for 2 minutes, agitating at least every 30 seconds. When compared to development, this etch time is significantly more robust to error. It
Figure B.4: Backside of an SOI wafer after etching exposed chrome and gold and removal of the photoresist mask.

is also independent of how many times the wafer is removed from the etchant. Nevertheless, significant over etching can reduce the size of the metallization by undercutting the photoresist. Once the 2 minutes has passed, remove the wafer and attempt to see if any unprotected gold features remains. This can be difficult to determined due to the dark opacity of the etchant, but large sections of gold can still be correctly identified. If gold remains, submerge again for another 2 minutes with agitation and repeat as necessary. Once gold appear to be etched, rinse with DI water in the sink, place in the DI water bath, and observe again. If gold remains, blow dry the wafer the with set procedure and submerge again in the gold etchant for another 2 minutes with agitation and repeat as necessary. It is important insure that water does not contaminate the gold etchant, for this can also reduce the etch time and may lead to early disposal of the liquid.

Once the gold features have been etched, chrome should be visible underneath. Deposited chrome looks very similar to silicon, but by flipping the wafer back and forth to compare it to actual silicon, it can be seen that the chrome has more of a grey color. The dry wafer can then be submerged into the bath of chrome etchant, following the same procedure as was outlined for gold. Etching chrome is a bit easier than gold because the solution is transparent. Upon close inspection during etching, the chrome can be seen to eventually disappear. Once this occurs, give the solution at least 15 more seconds to ensure it is completely gone, before removing it from the solution, rinsing, submerging and drying.

Once the metal has been successfully patterned, the masking photoresist must be stripped and the wafers cleaned. This is performed with a two step process: 1) Removal of a bulk of the photoresist using a solvent, and 2) Thoroughly cleaning with an aggressive cleaning solution. For the first step, acetone is usually the go-to solvent for aggressive room-temperature cleaning; however, the developer of the photoresist has been observed to produce faster results. Either chemical is acceptable, as this first step is simply to prevent unnecessary contamination of the second solution to produce a cleaner surface. Soak each wafer in either solvent for 10-15 minutes. Once complete, rinse with DI water.
in the sink and blow dry, using the procedure previously described.

For the aggressive cleaning solution, there are two options: 1) RCA-1, and 2) Piranha. In the past RCA-1 was used and it seemed to produce adequate cleaning. In attempts to continue to improve the cleaning process, RCA-1 was replaced with a 1:1 solution of Piranha (equal parts sulfuric acid and hydrogen peroxide). Piranha is significantly more dangerous to work with than RCA-1, but it produces a more aggressive cleaning solution. Each batch of RCA-1 or Piranha only lasts approximately 30 minutes. It is best to only attempt to clean three wafers with each batch, for 10 minutes per wafer. Once complete, rinse with DI water, submerge the wafer in a DI wafer bath, then blow dry, using the procedure previously described. Once complete, the backsides of the wafers should appear similar to Figure B.4. Figure B.5 shows the same result with a more standard backside metallization mask: 8 mm die sizes with 2 mm diameter gold spots in the center of each.

After this cleaning step, the wafers will be ready for front-side processing. Before loading these clean wafers back into their individual wafer carriers, it may be beneficial to clean the interiors of the carriers. This can be done by soaking a Texwipe with IPA and gently rubbing the interior of the case and spring. It is acceptable to place wafers with photoresist inside these carriers (and is required); however, cleaning them of any remnants of this photoresist at the same time the wafers themselves are cleaned is ideal.

### B.2.5 Front-side Processing (optional)

When wire bonding silicon devices, the industry standard is to never wire bond directly on silicon, but rather gold. It is believed that wire bonding on gold results is a stronger, more reliable bond, especially when considering thermal cycling. As experimentally observed in the laboratory, this is not the case. With the correct wire bonding parameters, an aluminum-to-silicon bond can be created with a bond strength that exceeds the tensile strength of the aluminum wire. Many such devices have also been repeatedly thermally...
cycled without loss of performance. That said, it is also entirely possible that the observed sample size in the laboratory is too low and with a higher population of devices, the benefit would become more clear. For this reason, no judgements are being made on the usefulness of gold bond pads, but rather simply stating they are not necessary for prototype devices.

In addition to this, the addition of front-side metallization can significantly complicate the fabrication process. First of all, the wafers from Table B.1 must be modified: all properties can remain the same, but the surface silicon dioxide is not necessary. This layer can be easily stripped using hydrofluoric acid (HF), but the initial deposition would have been an unnecessary cost. Second, the hard mask can no longer be thermally-grown silicon dioxide, but rather externally deposited silicon dioxide using chemical vapor deposition (CVD). There are many processing parameters that influence the properties of CVD oxide, but generally speaking, thermally-grown oxide is nearly always preferred. It has better adhesion to silicon and serves as a more resistant mask, when compared to CVD oxide. This means that a thicker layer of CVD oxide is required to mask the silicon device layer during etching, and depending on the properties of the CVD oxide, the increase in this thickness may vary depending on the deposition equipment. This is not to say that that CVD oxide is not useful, but there are simply more parameters to track. In the past, CVD oxide thickness between 2 and 3 $\mu m$ have successfully masked 100 $\mu m$ of silicon device etching, compared to 1.5 $\mu m$ of thermally-grown oxide. Third, because metallization has already been patterned on the backside of the wafer, using the metal etching technique on the front-side as well can lead to complications. Any method used to protect the backside of the wafer during etching will contaminate the front-side of the wafer; therefore, the second technique for depositing metal traces must be used: lithography before metal deposition.

The process begins by first performing lithography on the front-size of the wafer. This process is very similar to the previously described process: solvent rinse, dehydration, photoresist spinning, soft bake, exposure, and development. It is important that after developing, the cross-section of the patterned traces must be dovetailed, with more photoresist being removed at the silicon interface than at the air interface. To create this profile, a specific type of photoresist can be used: AZ nLoF 2035. AZ nLoF 2035 is a negative tone lift-off photoresist, while AZ4620 is positive tone. This is important to consider when designing masks for the process, as well as during deposition, which requires the use to two separate photoresist spinners. Please refer to the product’s datasheet when choosing spin and exposure settings.

With the photoresist applied, the wafers can then be loaded into the thermal evaporator and the same process followed as previously described. Once complete, the wafers will be coated with metal on the front-side, but the features should still be visible, almost resembling an embossing. To remove the metal that is sitting on top of the deposited photoresist, the wafers must then be soaked in either a photoresist stripper or acetone for a lengthy period of time. Generally, this requires at least overnight, but can take a few days, depending on the size of the features being removed. After the soak, removing and rinsing with a solvent through a squeeze bottle may help in removing tenacious metal flakes.
Once complete, the wafers must then be cleaned again using an aggressive high-temperature clean, such as RCA-1 or Piranha. The details of this process can be found in the previous section. The result are SOI wafers with patterned gold on both the front and backsides of the wafers. The front-side of such a wafer with metallized bond pads is shown in Figure B.6.

The front-side silicon dioxide hard mask must now be applied to the wafers. This process is typically outsourced to an external vendor and requires anywhere from 2 to 3 $\mu$m of CVD silicon dioxide to be deposited to successfully mask the silicon throughout the etching of a device layer with a thickness of 100 $\mu$m. The thickness of this CVD oxide varies depending on the deposition settings. Once the wafers return from deposition, the process merges once again with the standard procedure without front-side metallization, as continued in the following section.

B.2.6 Lithography on Silicon Dioxide

At this point, the backsides (and possibly front-sides) of the wafers have been completely processed and the entire wafers have been cleaned. Since this process, or possibly from simple exposure to the wafer carriers, particulates may have fallen on the wafers. To ensure cleanliness, another solvent rinse is performed immediately before processing. This rinse is conducted in a similar way as previously described: acetone, IPA, DI water, blow dry.

Upon completion of this process, the wafers must again be dehydrated at 120 °C for 1 hour. Up to this point, the lithography process has been identical as to when depositing on metal or silicon alone. Performing lithography on silicon dioxide adds additional difficulty because it is strongly hydrophilic. This property pulls humidity out of the air and absorbs it into the material, leading to poor photoresist adhesion. For this reason, dehydrating the wafers immediately before lithography can help promote adhesion by purging the wafer of any absorbed water. While this can aid in the initial adhesion of the photoresist to the wafer, the physical properties of the wafer remain the same. This can cause issues during development, the solution of which largely consists of water. Depending on the degree of hydrophobicity of the wafer, water in the developing solution can absorb into the exposed silicon, effectively undercutting the photoresist features.
### Table B.2: Properties of photoresist AZ4620.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Tolerance</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solids Content:</td>
<td>41.7</td>
<td>±1.4</td>
<td>%</td>
</tr>
<tr>
<td>Kinematic Viscosity @ 25°C:</td>
<td>360</td>
<td>±60</td>
<td>cSt</td>
</tr>
<tr>
<td>Absorptivity Specification @ 398nm:</td>
<td>1.00</td>
<td>±0.04</td>
<td>1/g cm</td>
</tr>
<tr>
<td>Specific Gravity @ 25°C:</td>
<td>1.07</td>
<td>±0.01</td>
<td></td>
</tr>
<tr>
<td>Water Content:</td>
<td></td>
<td></td>
<td>0.5% max</td>
</tr>
<tr>
<td>Principal Solvent:</td>
<td>2-ethoxyethyl acetate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appearance:</td>
<td>Clear Amber-Red</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Particle Count:</td>
<td>&lt;5</td>
<td></td>
<td>#/in²</td>
</tr>
<tr>
<td>Filtration:</td>
<td>0.5</td>
<td></td>
<td>µm</td>
</tr>
<tr>
<td>Flash Point:</td>
<td>98</td>
<td></td>
<td>°F</td>
</tr>
</tbody>
</table>

While this may not be a concern when patterning silicon alone, which has only a thin, native silicon dioxide film, thicker layers of silicon dioxide may still have issues.

For this reason, after dehydration, a film of HMDS ([(CH₃)₃Si]₂NH, or Hexamethyldisilazane) can be applied to the wafer to increase its hydrophobicity. This process is also known as ‘priming’ the wafer, and effectively silanizes the surface to make it hydrophobic. Wafers can be primed in this way through one of two common techniques: 1) Spinning the chemical onto a the wafer immediately before applying the photoresist, or 2) Allowing the chemical to coat the wafers by vaporizing the liquid in a contained environment. A thick coating of the chemical should be avoided because only a surface treatment is required to increase the hydrophobicity of the wafer; when thick coating are applied, there are additional issues with adhesion between the photoresist and HMDS, which can null any benefits of the process. For this reason, vaporizing the liquid is the strongly preferred method. This can easily be done by placing both the wafer and an open container of HMDS in a vacuum chamber and reducing the pressure to below 1 Torr. HMDS has a vapor pressure of 20 Torr, so this allows the chemical to vaporize and begin coating the entire interior of the vacuum chamber, including the wafers. In UCI’s INRF, the YES HMDS oven is dedicated to this process.

Once the wafers have been dehydrated and primed, AZ4620 photoresist can be applied to the wafers, using the same process as previously outlined, though with a modified development time. In the past, AZ4620 resist was used for lithography because it alone can be used as a mask for etching silicon in the STS DRIE in UCI’s INRF. Devices have been successfully produced using this process with silicon device layer thicknesses up to 50 µm and features sizes of 7 µm, producing aspect ratios of ~ 7. In order to pursue higher aspect ratios, the etching step was transferred to UCLA’s Plasma-Therm FDRIE, which was then shown to demonstrate the etching of silicon device layer thickness of up to 100 µm and aspect ratio of ~ 14. AZ4620 can not be used as a mask for etching in the FDRIE due to high temperatures during the etching process, therefore a silicon dioxide hard mask was incorporated. Currently, AZ4620 is still being used to mask wafers for transferring features to the hard mask; however, this is not ideal due to the thickness of the photoresist. The recipe for AZ4620 given previously produces photoresist thickness...
of $\sim 7 \, \mu m$; this thickness is significantly more than what is required to etch 1.5 to 3 $\mu m$ of silicon dioxide. It is recommended that a thinner photoresist is used in the future in order to improve the robustness of the lithography process by reducing dependance on exposure and development time on feature tolerance, as well as reduce development time to lengthen the window of acceptable development times.

Currently, development times of 1 minutes and 30 seconds have been used to successfully develop device features; though, this is largely influenced by the features of the mask. Because of how quickly the photoresist begins to debond during development, it is recommended to further modify the development time by increments of only 15 seconds until the correct time is found. As previously mentioned, as wafers are processed from any given lithography mask, one of the goals is to find the correct development time, requiring the wafer to only be submerged once. Submerging multiple times can lead to asymmetric development, leading to feature size variability across the wafer.

For reference, the properties of AZ4620 are given in the table below, as referenced from its published datasheet [213].

When examining the wafer to determine if development has been successful, it is important to first examine the large dicing trenches between sensors, as these regions are easiest to see, followed by the smallest features on the mask, such as comb teeth and gaps in the parallel plates. When the mask has not been developed long though, discoloration can be seen in the gaps. Figure B.7 displays a microscope image of a wafer after successful development, specifically focused on the parallel plate electrodes. Please note the difference in coloration between the features and gap.

Figure B.7 also shows the successful alignment of the sensor features and front-side metal bond pads. Should front-side metallization be a part of the process, it may be tempting to align the lithography mask to the front-side gold instead of the backside metal: this is not recommended. If possible, all subsequent masks in a multi-mask process
should be aligned to the first deposited mask. This avoids compounding alignment errors, and though not crucial for this specific process, it is good practice. In addition, backside alignment is significantly easier in the MA6, allowing the user a full range of view of both the mask and wafer. With front-side alignment, the user must look through opening in the mask to align.

B.2.7 Silicon Dioxide Hard-Mask Etching

With the front-side of the wafer patterned with device features in photoresist, the next step is to etch the silicon dioxide hard mask. This can be performed using at UCLA’s Nanolab, using the STS AOE (Advanced Oxide Etcher). The recipe that is used in this process is named “oxidapic” and the parameters are given in Table B.3.

To etch 1.5 µm of thermally-grown silicon dioxide, the total process time generally takes less than 2 minutes; however, etch rate continuously varies in this machine along with temperature due to high helium leak rates. Helium is used in the machine to cool the backside of the wafer during the etching process. If the surface of the wafer is contaminated, or previous users placed materials in the machine that were contaminated, a poor seal is made around the edge of the wafer, allowing helium to leak from the backside of the wafer to the front. To help prevent this, always clean the machine for 30 minutes prior to use using a clean dummy wafer and the “O2clean” recipe. This is a useful preventative measure, but will not solve the problem should their be major contamination of the machine, or if your wafers are contaminated. Please note that the seal is made on the backside of the wafer, so backside contamination is the issue. Performing an “O2clean” on a dummy wafer can also help determine where the contamination is located: If the helium leak rate is high using a clean dummy wafer, the issue is contamination of the machine. If the helium leak rate is low using a clean dummy wafer, but high on your wafers, the issue is the cleanliness of the backside of your wafers. For this reason, the helium leak rate should always be recorded for each wafer.

Due to the variable etch rate over time and between wafers, the proper etch time should be calculated for each wafer. During the 30 min O2 cleaning of the machine, the initial oxide thickness of the wafers can be measured using the Nanospec, which is located near the STS AOE. This machine should be used to measure and record the oxide thickness at five locations across each wafer: center, top, bottom, left, and right. This is accomplished by choosing the highest objective lens of the microscope (40x) and focusing it on the silicon between the dicing strips in each general location. There should be no visible photoresist in the image for an accurate measurement.

Once complete, the wafer can then be etched for between 30 and 60 seconds; 60 seconds if the helium leak rate is low, or 30 seconds if it is high. When the helium leak rate is high, the temperature of the wafer increases and can burn the photoresist on the surface, potentially causing reflow which can deform the features. Reducing the etching time to 30 seconds reduces the maximum temperature reached inside the machine, which aids in mitigating this issue. If the reason for the high leak rate is due to contamination of the machine, it may also help to rotate the wafer in the chuck in 90 degree increments for each etching iteration. This can aid in achieving etching symmetry. Please note that minimally, each wafer will need to be loaded into the machine twice: once to etch the
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Process Times</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pump Down Time:</td>
<td>30</td>
<td>sec</td>
</tr>
<tr>
<td>Gas Stabilization:</td>
<td>15</td>
<td>sec</td>
</tr>
<tr>
<td>Process Time:</td>
<td>30-60</td>
<td>sec</td>
</tr>
<tr>
<td>Pump Out Time:</td>
<td>10</td>
<td>sec</td>
</tr>
<tr>
<td>Recipe Process Mode:</td>
<td>Discrete</td>
<td></td>
</tr>
<tr>
<td><strong>Pressures</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base Pressure:</td>
<td>0.1</td>
<td>mTorr</td>
</tr>
<tr>
<td>Pressure Trip:</td>
<td>94</td>
<td>mTorr</td>
</tr>
<tr>
<td><strong>Gas Flow</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_4F_8$:</td>
<td>7</td>
<td>sccm</td>
</tr>
<tr>
<td>$Ar$:</td>
<td>10</td>
<td>sccm</td>
</tr>
<tr>
<td>$CHF_3$:</td>
<td>33</td>
<td>sccm</td>
</tr>
<tr>
<td><strong>Power Generation to Coil (13.56 MHz)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Power:</td>
<td>700</td>
<td></td>
</tr>
<tr>
<td>Matching:</td>
<td>Auto</td>
<td></td>
</tr>
<tr>
<td>Match Load:</td>
<td>65</td>
<td>%</td>
</tr>
<tr>
<td>Match Tune:</td>
<td>32</td>
<td>%</td>
</tr>
<tr>
<td><strong>Power Generation to Platen (13.56 MHz)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Power:</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>Matching:</td>
<td>Auto</td>
<td></td>
</tr>
<tr>
<td>Match Load:</td>
<td>40</td>
<td>%</td>
</tr>
<tr>
<td>Match Tune:</td>
<td>41</td>
<td>%</td>
</tr>
<tr>
<td><strong>Helium Flow</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBC Active:</td>
<td>On</td>
<td></td>
</tr>
<tr>
<td>Pressure:</td>
<td>9000</td>
<td>Torr</td>
</tr>
<tr>
<td>Max Flow:</td>
<td>40</td>
<td>sccm</td>
</tr>
<tr>
<td>Min Flow:</td>
<td>10</td>
<td>sccm</td>
</tr>
</tbody>
</table>

Table B.3: Silicon dioxide etching recipe for the STS Advanced Oxide Etcher.
Figure B.8: Frontside of an SOI wafer after etching of the silicon dioxide hard-mask: the full wafer (left), along with a focused view on a single device (right).

Figure B.9: Frontside of an SOI wafer after etching of the silicon dioxide hard-mask: portion of a device (left), along with a closer view of the smallest features (right).
wafer incompletely for the purpose of calculating etch rate, and a second time with the updated etch rate to determine the correct etching time.

With each wafer etched at least once, the oxide can once again be measured for each wafer in the same five general locations: center, top, bottom, left, and right. To ensure complete etching, the slowest etch time should be calculated for each wafer by taking the lowest value observed in the initial oxide thickness and subtracting the highest value observed in the current oxide thickness. Using this value to calculate etch rate, it should then be applied to the highest value observed in the current oxide thickness to produce the remaining etching time. This time will be different for each wafer, but it is usually no more than 10 seconds if the wafers are clean.

Once etching is complete, the photoresist mask must be removed. This can be accomplished using the Matrix asher in UCLA’s Nanolab. The recipe used is named “3 min STRIP”, with properties recorded in Table B.4. The process time can not be modified on this machine, therefore to ensure complete removal of the photoresist mask, each wafer must be run 4 times, for a total of 12 minutes for each wafer.

The lithographic pattern has now been successfully transferred to the silicon dioxide, along with the removal of the organic photoresist mask. The wafers should now resemble that of Figures B.8 and B.9. Figure B.8 displays images on both the wafer and device scale, while Figure B.9 shows microscope images using two different lens objectives.

### B.2.8 Silicon Device Etching

The silicon wafers are now free of organic residues and their device layer is masked with silicon dioxide. Etching of the device layer can now be performed in UCLA’s Plasma-Therm FDRIE. When using this machine, there are a few issues to keep in mind. First, wafers with an unpatterned, blanket coating of chrome and gold on the backside of the wafer had a high probability of not being able to be unloaded from the chamber after etching. It is believed this was due to the smoothness of the metal coating reducing the friction between the wafer and the unloading equipment. This issue has since been solved by patterning the backside metallization; however, another method that was used was to apply plastic dicing tape to the backside of the wafer, and trim the parameter using a razor blade. The tape must be able to resist the high temperatures of the etching process, but this method was also successful at increasing the friction on the backside of

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pressure</td>
<td>3.75</td>
<td>Torr</td>
</tr>
<tr>
<td>MFC1</td>
<td>45%</td>
<td></td>
</tr>
<tr>
<td>MFC2</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>MFC3</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>RF Forward</td>
<td>500 W</td>
<td></td>
</tr>
<tr>
<td>Pins</td>
<td>Down</td>
<td></td>
</tr>
<tr>
<td>Temperature</td>
<td>250°C</td>
<td></td>
</tr>
</tbody>
</table>

Table B.4: Photoresist stripping recipe for the Matrix System One Stripper.
the wafer, and may be a useful trick in the future.

Second, like the STS AOE, the etch rate in the FDRIE can drift over time. This can make it a challenge when attempting to avoid over-etching a wafer due to ‘footing’ effects, as will be later explained. For this reason, it is generally desired to etch each wafer in a batch with a different etch time in an attempt that at least one wafer will be etched with an optimal time. The dispersion range of the etch times should be a function of the length of time that has passed since last using the machine, and hence the etch rate confidence interval. In the past, typical batch sizes have been of three wafers, where one was etched with the anticipated etch time, and the other two wafers were etched with times 30 seconds longer and shorter. Please note that in the past, etching a silicon device layer with a thickness of 100 $\mu m$ and minimum gap of 7 $\mu m$ would require anywhere from 30 to 35 minutes, using the recipe “DarkField2”. For future reference, the parameters of this etching recipe are supplied in Table B.5.

One of the main benefits of using the FDRIE is for both a quicker etch rate, as well as minimal sidewall scalloping through faster transitions between the etching and passivation sequences. Sidewall scalloping is an unfortunate side-effect of DRIE, created due to the discrete series of steps that are cycled throughout the process. A visual representation of this effect is shown in Figure B.10, which also demonstrates the DRIE etching process. This scalloping can lead to spring / mass asymmetry which can deteriorate device performance.

Additionally, when using a dielectric as an etch stop, such as the buried layer of oxide in SOI wafers, continued etching once the dielectric has been reached can lead to lateral silicon etching at the silicon-oxide interface, as shown in Figure B.11. This occurs due to the ion bombardment during the RIE process: the charged ions are forced into the dielectric, however, the material quickly becomes charged due to its insulative properties. This charge then repels any further ions from contacting the material, redirecting them into the sidewall and etching the passivation layer at the interface. Like sidewall scalloping, this can also lead to spring / mass asymmetry.

Table B.5: Deep silicon etching recipe for the Plasma-Therm FDRIE.
Figure B.10: DRIE process flow, used for deep etching of silicon with vertical sidewalls.

Figure B.11: DRIE when a buried oxide layer is reached. Charging of oxide during RIE etches the sidewall passivation, creating excessive silicon at the interface.
Figure B.12: Frontside of an SOI wafer after etching of the silicon device layer: the full wafer (left), along with a focused view on a single device (right).

Figure B.13: Frontside of an SOI wafer after etching of the silicon device layer: portion of a device (left), along with a closer view of the smallest features (right).

Figure B.14: Frontside of an SOI wafer after etching of the silicon device layer: view of the proof mass with etch holes (left), along with a closer view (right).
While it may seem the answer is to simply not over-etch to prevent this from occurring, this is easier said than done. Any variation in trench width influences etch rate, and variability in gap width is required for multiple rows of capacitive electrodes. In addition to this, it is not just trench width that influences etch rate, but also any intricate design elements. For instance, trenches that merge in a “L”, “T”, or “X” joint have a faster etch rate. Abrupt change in width may also aid in increasing the etch rate of the narrower gaps. As an example, consider the comb drives of Figure B.13. The gap between the comb fingers is 5 µm; however, due to the close proximity to the larger 15 µm trenches on either end, this 5 µm gap has a faster etch rate than long 7 µm trenches, which also appear on this device. While this effect can not be completely eliminated through sensor design, it can be minimized. Spring asymmetry is several orders of magnitude more critical than mass asymmetry, in terms of ‘footing’ effects [156]. For this reason, the trenches around the springs should be the thinnest gaps and the last to etch through. Due to the benefits from increased capacitance, this minimum gap should be identical to the parallel plate gaps as well. By then attempting to achieve the optimal etch time for these gaps, the critical features of the device can be optimized while sacrificing a degree of mass asymmetry.

Once etching of the device is complete, it is recommended to clean the wafer once again using O2 plasma in the Matrix asher. There are two reasons for this. First, after etching, there is still a passivation layer on the etched structure which can only be removed with a high-temperature process. It has been hypothesized that leaving this coating on the silicon structure can lead to increased damping, either through larger thermoelastic damping or imparting additional stress to the structure. Second, depending on prior etching processes in the equipment, the etching chamber may be contaminated with unknown materials. These unknown materials have been known to redeposit on subsequent wafers, so removing these materials after etching is desirable for consistency.

The wafers should be cleaned with the identical process that was used to remove the photoresist mask after hard mask etching. To reiterate, on the Matrix asher, this involves running the “3 min STRIP” recipe a total of 4 times on each wafer, constituting 12 minutes of etch time.

Optical photographs of the wafers after etching are shown in Figures B.12 through B.14. There are a few things to note in these photographs. First, notice the darkened outlines of the silicon dioxide mask where it boarders the trenches. This is due to ion bombardment stripping the lateral passivation layer from the surface of the hard mask, resulting in increased mask deterioration at the trench edges. This is not a critical effect for this process, but could be with additional process development, so is worthy of note. Second, close inspection of the wafers, Figures B.13 (right) and B.14, reveals dark trenches, while in a broader view, Figure B.13 (left), they begin to lighten. This is due to inadequate light reaching the bottom of the trench, caused by the high aspect ratio cavities and short distance to the objective lens. Both factors block light from reaching the examined area. From experience, it is impossible to view the bottom of trenches with widths below ~ 20 µm using high objective lenses (40x), while for moderate objective lenses (10x) it is possible. This can be an important factor when attempting to view contaminates that may have lodged in certain features, as well as identifying etch depth.
to avoid over-etching. As was shown in Figure 5.12 of Chapter 5, optical profilometry can be used to monitor high aspect ratio (∼ 15) etch progression; however, this technique is limited by the 2-D resolution of the profilometer, as well as the absolute trench-width. While the aspect ratio of the device from Figure 5.12 is identical for the device of Figures B.12 through B.14, the trench width was reduced from 20 \( \mu m \) to 7 \( \mu m \), preventing sufficient light for an accurate measurement.

Other non-destructive methods exist for assessing etch rate as well. For instance, when etching a conductive device layer, electrical probing of two regions that will become isolated once etching is complete can be a binary indicator. This method could also be used to monitor etch progressive through careful assessment of changes in resistance. Another example is through infrared microscopy. There is currently some evidence to suggest that the angled edges that form at the bottom of etched cavities are capable of scattering light. By using a wavelength of light that is transparent through silicon, light that passes through these locations scatters, creating dark regions when detected on the other side of the device. An example of this can be seen in Figure B.15 (left).

The silicon device layer of this sensor was etched incompletely, leaving dark regions at the etch holes of the proof mass and narrow gaps of the capacitors and springs. Larger regions, however, such as the electrode anti-gaps have etched through to the smooth buried oxide. The device was then diced and a cross-section at the etch holes of the proof mass were examined using a scanning electron microscope (SEM), shown in Figure B.15 (right). This image clearly shows incomplete etching of the device layer (bottom), with the etched features never reaching the buried oxide (thin white strip).

IR images of devices after incremental etching are shown in Figures B.16 and B.17. After 1.75 min of additional DRIE etching, Figure B.16 (left), the proof mass etch holes have opened; however, the springs and narrow gaps of the parallel plate electrodes still have silicon that remains. Please note that the trench width of these latter two features are identical: 7 \( \mu m \). Regardless, the edges of the electrode gaps has already begun to break through due to the edge effects created by the large anti-gap of the feature. In
comparison, only the tip of the springs has etched through. At 2.0 min the springs begin to open, Figure B.16 (right); however, asymmetrically, revealing a bias during DRIE etching which is dependent upon not only feature size, but orientation. At 2.25 min the springs and electrodes have completed opened to the buried oxide, Figure B.17 (left); however, should etching continue, footing effects soon become apparent across the wafer, Figure B.17 (right). This technique can be used as a binary indicator of the DRIE etching process; however, please note that uniform etching across the wafer requires significant process development. Because such development is difficult at shared, academic facilities, a second option is to etch a majority of the device layer at once, then etching individual dies after dicing. In this way, the total etch time can be modified for each die across the wafer, leading to optimal etching without additional recipe development. Concerning the FDIRE at UCLA, this can be accomplish by simply using a silicon dioxide coated handle wafer, and placing the dies on top without need for attachment. This can be a useful
Figure B.18: SEM of a cross-section of the device, located at the proof mass, for two different etching times.

Figure B.19: SEM of a cross-section of the device, located at the parallel plate electrodes, for two different etching times.
Of course, destructive methods also exist for assessing etch rate, the most common of these being dicing and imaging using a scanning electron microscope. This technique can provide valuable information on etch rate and quality through the use of the high resolution images that can be obtained. Figures B.18 through B.20 display cross-sections of three distinct features of a quadruple mass gyroscope for two different etching times: the proof mass etch holes, parallel plate electrodes, and comb drives, for etch times of 33 and 32.5 minutes. The 'footing' effects due to over-etching can be clearly seen in these images, along with the lateral evolution of this effect over time, as shown in the insets of Figures B.18 and B.19. It is also worthy to note that this lateral distance is different for each feature, due to the variable etch rate between these regions, caused by the feature dimensions.

Despite the detailed information that can be extracted from these images, they unfortunately sacrifice the wafer, which prevents further etching should the buried oxide have not been reached in each location. This presents a need for fabrication methods which are robust to this kind of error, or at least the development of non-destructive techniques of assessing etch depth on the wafer-scale.

**B.2.9 Wafer Dicing**

After the silicon device layers of the wafers have been successfully etched, the wafers must then be diced to separate each die for packaging. Conventionally, this has been performed using a dicing saw; however, laser dicing has recently become a popular alternative.
Figure B.21: Frontside of an SOI wafer after device etching and sensor protection with photoresist, prior to dicing.

Dicing saws literally consist of a rotating blade which mechanically removes silicon in straight lines with precision motion, while laser dicing uses a point source which can potentially be used to dice intricate shapes. Furthermore, mechanical dicing requires lubrication, which typically floods the dies in liquid, potentially contaminating the surface with external materials or silicon debris, while laser dicing is a completely dry process with no concern for contamination. It is also important to note that the costs of these two processes, as charged from external vendors, is comparable. It is believe that the only complication of laser dicing is that it can not be used to cut through metal films, though at the time of this writing, this method has not been tested. For this reason, the continued description of the fabrication process will assume the use of a dicing saw for die singulation.

To protect the wafer during dicing, a coating of photoresist must be applied to the wafer. The same photoresist used for sensor definition can be used, AZ4620, though for a thicker coating it is recommended to reduce the spin speed form 5000 $rpm$ to 2000 $rpm$. The soft bake time should also be reduced from 20 $min$ to 3 $min$ to make the removal of the resist easier. Finally, after the soft bake, exposure is not required. Simply load the wafers back into their individual wafer carriers. A photograph of an etched silicon wafer with a protective coating of photoresist is shown in Figure B.21.

In the past, dicing has typically been preformed by outside vendors to minimize chipping. Several different vendors have been shown to be capable of excellent work, with contact information listed in Appendix D. To avoid confusion, in addition to mailing the wafers, emailing a description of the wafers along with a dicing diagram is recommended. The description should include detailed information on the material and thickness of each layer, while the image should show the specific grid to be diced overlapping the wafer. Like most machining processes, for each cut axis only one alignment is made with subsequent cuts made in relation to this first cut. Keep this in mind when dimensioning the image. These cuts will generally form a grid to isolate each die, plus any additional cuts to create cross-sections for SEM imaging.
Depending on the priority of the order, the wafers will return after several business days with all necessary cuts, held in place with dicing tape. As a preliminary measure to remove any particles from the surface of the dies, gently blow dry the surface of the wafer in an INRF fume hood using nitrogen at a high angle of attack. Once complete, proceed to removing a batch of sensors (8 to 12) from the tape for processing. To remove from the dicing tape, hold the tape upside-down over a Texwipe. Avoid allowing the dies to push against the Texwipe, as this can also be a source of contamination. With tweezers, gently push on the back of the dicing tape beneath the sensor that is to be removed. They may require repeatedly moving across the end of the tweezers across the die, but eventually the die will fall onto the Texwipe. Repeat this procedure until each required die is removed, and return the remaining dies on the tape to their protective holder for storage.

The first step to cleaning the dies is the removal of a bulk of the photoresist through the use of a solvent rinse. This is completed using two separate baths of acetone, and a third of DI water. Begin by holding each die upside-down in the first bath of acetone, suspending in the liquid with the tweezers. Hold the die like this for 15 seconds before removing the die and placing it right-side up in the second acetone bath. This procedure attempts to remove any particles left on the surface of the photoresist from the dicing process through gravity. The acetone quickly attacks the surface of the photoresist, undercutting any particles and pulling them away from the surface of the die to the bottom of the acetone bath. The dies are then allowed to continue to soak in a second, clean acetone bath for further cleaning of stubborn resist that may have flowed into the trenches. Once each die has been processed in this way and is sitting in the second bath of acetone, attach a lid and allow the dies to soak for approximately 10 minutes. Once complete, transition the dies to the bath of DI water to serve as a buffer before transition to a more aggressive cleaning solution. The two solvent baths may now be disposed.

For a more aggressive clean to ensure the removal of the photoresist, RCA-1 or Piranha must be used. As previously mentioned, a 1:1 solution of Piranha is preferred; however, requires the handling of an experienced user. Please note that cleaning dies is further complicated by the DI water buffer. Just as sulfuric acid reacts with hydrogen peroxide, it can also react with just water. For this reason, attempt the gently shake off any water from each die before transferring to the Piranha solution; even despite this, the reaction will generally become more aggressive as additional dies are added. Due to the time required to transition the dies in or out of the liquid (2 min from personal experience), to ensure that each die experiences the same amount of cleaning time, it can be useful to line the dies along the rim of the container, clockwise, when adding them to the solution. When the cleaning time is complete, the dies can then be removed from the solution, clockwise once again. For cleaning, this is not as critical, but can be a useful technique when etching a number of discrete components. Please note that not observing any safety protocol instated by the INRF is not recommended. With that stipulation, please note that the thick gloves required when working with acids significantly reduces manual dexterity, which can potentially result in damaged to dies by dropping to gouging with the tweezers. Please keep this in consideration.

Once the high temperature cleaning is complete, the dies should be transferred again
Figure B.22: Clean silicon dies with surface silicon dioxide, ready for hydrofluoric acid release.

to a new container of DI water (the previous container should be disposed, triple rinses, and refilled). This solution dilutes any residual acid from the dies through diffusion, and should be left in this container for at least 2 min. During this time, a second bath of IPA should be prepared, and once the DI water soak is complete, the dies transferred. Once again, the dies should be left in this container for at least 2 min. Soaking the dies in IPA is useful for speeding the drying process, as IPA evaporates much faster than water. IPA to chosen for this process instead of acetone due to higher chemical purity, therefore fewer contaminates.

Once complete, one of two techniques can be used to dry the dies, depending on the user: 1) Evaporation, or 2) Blow drying. The use of each technique is debatable, therefore both will be presented. Drying by evaporation involves simply removing the dies from IPA, gently shaking to remove any drops, then gently placing on a Texwipe, face-up. The dies are lined up and covered with a Petri dish, which is propped up on one side to allow air flow. After a few minutes, each die is then moved to a new location on the Texwipe, recovered, and allowed to continue to evaporate. While the dies should never be placed upside-down on a Texwipe after cleaning, the purpose of placing them at a new location is to aid in the drying of the backsides of the dies. An alternative approach to this method also involves placing the dies on a hotplate set to a moderate temperature to speed evaporation. Once complete, the dies can be transferred to a Petri dish. Please note that is it very important that the backside of the dies are dry, otherwise they can stick to the dish, making it very difficult to remove, which usually results in destroying the device.

The second approach is blow drying. Upon removal of each die from IPA, continue to grip the die and hold in contact with a Texwipe with one hand. With your second hand, gently blow dry with nitrogen at a high angle of attack. Transfer to a new location on the Texwipe, still face-up (to dry the backside) and blow dry again. Repeat until both sides appear dry and place in a Petri dish and repeat until all dies are dry. The blow drying method is believed to result in lower contamination because less IPA is allowed to dry on
the dies. However, it may also result in increased contamination because more particles may be introduces from the tweezers, or simply the fume hood itself due to the strong air currents introduced by the blow drying. In addition, there is also the increased risk that a die may slip from the grasp of the tweezers and fly to the back of the fume hood, which generally destroys the device, or at least contaminates it. Due to the debatable benefit and potential issues with the blow drying method, the evaporation method is recommended.

Once dry, the dies are now clean and ready for release by hydrofluoric acid (HF). A photograph of a number of dies at this stage are shown in Figure B.22. Please note the various colors due to the silicon dioxide hard mask on the surface of each. Each die is at a slightly different angle, resulting in a variability of light refraction.

### B.2.10 Hydrofluoric Acid Release

The final step in the fabrication process, as outlined in Figure B.1, is the isotropic etch of the buried silicon dioxide. To accomplish this, one of two techniques can be used: 1) Liquid release, or 2) Vapor release. Liquid release involved soaking the dies for ~70 min in 20% HF, followed by a DI water and IPA soak, as mentioned in the previous section. Depending on the stiffness and gaps of the released structure, critical point drying can also be used to help prevent stiction. Figure B.23 displays an optical image of a device where the parallel plates stuck together during release, preventing future movement of the resonant structure. When this occurs, the device is generally destroyed. While liquid release can be successfully used for certain designs, the vapor-phase HF method nearly supplants it in every way.

Vapor-phase etching simply requires a small enclosed cavity, a reservoir of 49% HF, and a method of suspending the structure to be etched upside-down with temperature control. Commercial equipment exists to complete this process, as shown in Figure B.24. HF, along with water, naturally evaporates at room temperature and atmospheric pressure. Placing an exposed reservoir of HF in an enclosed space under these conditions
will result in condensation on the walls of the cavity. This is not ideal when attempting to etch a device, because the surface tension of the liquid is to be avoided to prevent stiction. To prevent the formation of droplets, the etched material must be elevated in temperature. While there is a certain critical temperature that must be applied to the substrate to prevent droplet formation, this temperature can also be used to modulate etch rate by controlling the ratio of HF to water vapor at the surface of the device. As the applied temperature rises, the ratio of HF to water vapor increases, resulting in a lower etch rate. While this may seem counter-intuitive, it is because water is a necessary ingredient for the etching of silicon dioxide with HF. By forcing water to be the limiting reagent, the etch rate can be modulated as desired [214]. Reducing etch rate can be desirable for achieving a precise silicon dioxide etch depth, or for increasing uniformity across a wafer. For this specific application, the uniformity achieved at higher etch rates is acceptable, therefore minimal device temperature is desired. Typically, 45 min of etching at 36 °C has been an acceptable procedure, though this etch time may need to be increased as the HF inside the equipment ages. Initially, 49 % HF was loaded in the machine, but over time, this concentration may change.

Once etching is complete, or even at intermittent intervals throughout the etch process, the silicon dioxide removal process can be monitored using an infrared (IR) microscope. Silicon is ∼ 50 % transparent to wavelengths of about 1 to 8 µm, as well as wavelengths throughout the 50 to 100+ µm range. Silicon dioxide also has a similar range of transparency. By using a microscope that produces and detects this light, the etch rate of the buried oxide layer can be monitored. This is accomplished using an Idonus IR-Light Microscope, shown in Figure B.25. This microscope functions similar to X-ray imagery: the light source and detector are on opposites sides of the object being imaged, for the purpose of observing the contrasts in transparency of the analyzed object. Bright regions indicate a level of IR transparency and digital filtering to saturate the image, while dark regions indicate the opposite effect. Dark regions can be created by poor material transparency, angled surfaces causing reflections, or simply digital filtering.

Examples of the IR imagery that can be obtained from this process are shown in Figure B.26, which displays the same area of a silicon die at three different times: 1,
2, and 3 hours of etching. The images are taken at the parallel plate electrodes of the structure and show the buried silicon dioxide progressively being removes as the length of HF exposure increases; using standard visual inspection, each of these photographs would look identical.

The microscope settings to obtain this imagery can vary greatly, depending upon both the material being examined, as well as the age of the light source. There are two pieces of information to draw from this: 1) It is critical to turn off the light source when not in immediate use to preserve its life, and 2) The useful information is obtained in the relative contrast of the image, not the absolute contrast. As a rule of thumb when imaging SOI wafers, the best images are taken when the physical IR light source level is maximized, with moderate levels of digital gain, gamma, and exposure. The exact levels of the digital parameters always require adjustment for the best results.

From experience, when silicon is being examined, the level of doping can influence transparency, with higher levels of doping leading to worse transmission, and less observability. The true relationship between silicon doping and IR transparency is currently unknown, but significant IR transparency variation can be associated with the doping tolerance in the SOI wafers of Table B.1. With higher device layer resistivity of 0.003 Ω·cm, excellent transparency is observed; however, with the lower resistivity of 0.001 Ω·cm, the buried oxide is difficult to observe. While every 4 inch wafer imaged from Ultrasil has been acceptably transparent to IR, 8 inch wafers experienced extreme difficulty, and
were experimentally shown to have lower resistivity.

While not explicitly required, the silicon dioxide etch rate can also be calculated. This is useful for monitoring the etch rate of the machine over time, or choosing an etch time for a new device design. The etch rate can be assessed by etching a test structure for a series of identical etch times, with removing and photographing the structure in the same location between each time increment. A scale bar for the images can be created by using the dimensions of the silicon device features. For additional accuracy, use a periodic feature, such as etch holes, and measure as many full periods as possible. Determine the number of image pixels and divide by the physical length of each period, as defined by the lithographic mask, multiplied by the number of cycles. By measuring a periodic feature, gap biases can be compensated, while a long length reduces the quantization error due to the discrete pixel count. Using the calculated scaling, measure the undercut distance of the silicon dioxide beneath the anchors in pixels and convert to a physical length. An example plot displaying the anchor etch depth versus time is shown in Figure B.27 for two different fabrication processes: the standard in-house SOI process (blue), along with the commercial SOI-MUMPS process (red). The variation in etch rate is attributed to the minimum gap sizes of the processes and thickness of the buried silicon dioxide layer.

Considering standard SOI devices, there is a significant window of acceptable etch times: the time must be long enough to fully release the structure, but short enough not to significantly compromise the structural integrity of the device and electrode anchors. The former can be confirmed by physically probing the device, while the latter can not be tested until the device is wire bonded after packaging.

Device probing is traditionally performed at a microscope probe station to allow precision motion under a microscope; however, it is easy to damage devices using this method by accidentally manipulating the z-axis control in the wrong direction. For large, high displacement silicon resonators, the wire bonding station can also be used. Simply position the die in a Petri dish under the station, allow approximately 5 to 10 mm of wire to extend from the tool, and use this wire to gently press on the proof mass from the side using an etch hole. This method requires a steady hand; however, should the
Figure B.27: Etch rate of buried silicon dioxide for devices from two different fabrication processes: in-house SOI fabrication (blue) and commercial SOI-MUMPS (red).

device be overextended, the additional wire will bend to alleviation that force before transferring to the structure. The larger the extended wire, the less maximum force that can be transferred.

Should the buried silicon dioxide layer be over-etched, the effect of this will not become apparent until after packaging and the device is being wire bonded. When attempting to attach a wire to the anchor, the force transmitted by the ultrasonic vibration will shatter the remaining oxide pillar and the entire support may be lifted into the air when attempting to draw out additional wire. It is best to find the correct etch time for any device to avoid this from occurring, due to the waisted processing time and materials.

Finally, once the dies have been fully released with vapor HF, it is recommended to clean them with O2 plasma. This appears to be the only negative consequence of vapor HF release; liquid HF release has never required this step. It is believed that contaminates either from the chamber, or within the silicon dioxide itself, and deposit on the surface of the silicon device during vapor release, while during liquid release they simply dissolve into the water. This is a speculation due to occasional issues that have arisen after vapor release and packaging, where wire bonding the structures proves to be difficult or impossible due to poor aluminum / silicon bonding. This have never occurred with structures that had been processed with a liquid release, or if an O2 plasma clean is conduced after vapor release. This difficulty is not always encountered if the final cleaning step is ignored, but it is a possibility.

B.2.11 Die Attachment

The individual silicon dies must now be packaged for electrical interaction, as well as protection from the environment. The first step to this process is the attachment of the silicon die to a ceramic package with conductive feed-through paths for electrical signals.
For a low-outgassing, rigid attachment, a gold / tin eutectic bonding step is completed between gold on the package, to the patterned backside gold of the silicon die. This requires that the two surfaces be brought into contact with one another, separated by a thin film of solder composed of 80% gold and 20% tin. A force must then be applied between the two components and the temperature elevated past the eutectic point of the material of 280°C. This allows the solder to reflow and bond with the gold on the two surfaces, allowing a rigid attachment that is stronger than the silicon itself.

During the process, and depending on the size of the solder preform and width of the patterned metallization on the backside of the die, the solder may reflow to regions of bare silicon. These regions will still bond to the package, but the strength of the bond will be significantly less. The reason for this weak bond is that gold can diffuse into silicon at elevated temperatures, and the solder itself consists of 80% gold. While the relatively low temperature of the die attachment process does not produce a significant level of diffusion, it is enough to create a bond that can impart stress to the device. For this reason, limiting the size of the solder preform is required.

To complete the die attachment step, a vacuum furnace is used, as shown in Figure B.28 with supporting equipment. The furnace is a UniTemp RSS-110 mini reflow solder system, along with supporting controller and power supply, interfacing PC, a dedicated vacuum pump, and chiller to protect the electronics in the system from the elevated temperatures. There are existing controller programs that have been saved on the PC for both die attachment and vacuum sealing. While these processes were initially calibrated for the dual in-line packages (DIP 24), as demonstrated visually in Figures B.29 through B.36, please note that it has also been successfully used for the leadless chip carriers (LCC 44) by simply swapping the two packages.

The first step of the process is to bake out the chamber, especially if the equipment
has not been used recently and exposed to ambient conditions. This step helps to ensure a clean working environment for the subsequent processing and allows a lower vacuum level to be reached by removing absorbed contaminates. This is completed by first wiping the interior of the chamber and graphite beams with IPA-soaked Texwipes. It is normal for the Texwipes to darken when wiping the graphite beams. Wait for the IPA to evaporate before loading the beams back onto the hotplate and beginning the bake out procedure. This procedure expose the chamber to vacuum, then heats the internal hotplate to the maximum temperature of 450°$^\circ$C for 1 hr.

As with every process, it is vital for the safety of the vacuum pump that the pump not be activated until the process has begun and the solenoid valve between the pump and the chamber is opened. The attached vacuum pump is a combination of a roughing the turbo pump, where the turbo pump is not activated until a set pressure level is reached. If the vacuum pump is turned on when this valve is closed, the turbo pump will activate as the pressure in the line is reduced. At this point, should the valve be opened, a significant amount of air would rush into, and potentially damage, the turbo pump. The same is true when each process is finished: the first step is to deactivate the vacuum pump and wait for the line to automatically purge before attempting any subsequent step.

With the chamber bake out complete, the next step is to bake out the packages and dies. Because both components are stored at atmospheric pressure, ambient molecules (primarily hydrogen and water vapor) absorb into the gold metallization where the primary die attachment is to occur. Baking out the metallization allows these molecules to release. Should this step not be completed, these molecules would release during the attachment and prevent a quality bond from forming.

If using DIP packages, the first step is to trim the leads of the packages. It is important to do this before die attachment to minimize the risk of sensor damage or contamination, and can be completed using any kind of equipment designed to cut thin metal, such as wire cuttings or tin snips. Packages before and after this procedure are shown in Figure B.29. The packages must then be loaded onto the graphite beams so that good heat conduction can take place directly to the backside of the packages. The released
silicon sensors should also be loaded into the chamber to bake out the backside gold. To allow ample exposure of the backside gold to the ambient vacuum, while still leaving the sensors upright to prevent front-side damage or contamination, each die can be propped up on one side using the graphite die attachment locators, as shown in Figure B.30.

The setup of Figure B.30 must then be exposed to vacuum and elevated in temperature to $320^\circ C$ for 2 hr. This temperature and time must be high and long enough to completely desorb any molecules from the metallization, yet low and short enough not to significantly age the packages. The package metallization consist of a gold plated nickel, and should the material be substantially heated for a long period of time, the nickel and gold can diffuse together into one solid alloy as opposed to two discrete metal layers. Should this diffusion reach a significant level, it can even be visually identified through a color change of the metallization: pure gold has a bright yellow appearance, while after significant diffusion the metallization develops a dully orange color. This type of aging is challenging because it reduces the purity of the surface gold, reducing the yield of both wire bonding and vacuum sealing. The rate of this diffusion is directly correlated to temperature. While literature supports a vacuum bake out at $400^\circ C$ for 1 hr [215], this temperature was reduced in an effort to reduce package aging to improve the yield of vacuum sealing. The process described in literature involved only a single bake out of the package, while the desired process involves two: once prior to die attachment and again before vacuum sealing. The chosen process has also been shown to produce void-free die attachment, as experimentally verified through X-ray imaging, as shown in Figure 6.6. While extensive fine tuning of this process has not been completed, please note that further reducing the time from 2 to 1 hr results in die attachment voids.

With the bake out complete, the packages and sensors must be assembled for die attachment. This is completed by unloading the samples from the vacuum furnace and first inspecting the dies under a microscope. In order to preserve the knowledge of the initial device orientation on the wafer, as well as the wire bonding schematic for asymmetric designs, the die code can be identified under magnification and used for orientation during die attachment. The assembly can then be formed by first placing in
the graphite locator within the package cavity. This locator is used to both center the
die within the package, as well as align the lid. A small piece of solder is then placed in
the middle of the package, which consists of an 80% gold / 20% tin alloy of dimensions
0.080 \times 0.080 \times 0.002 \, \text{in}. A smaller solder area can be used; however, the misalignment
error is high due to hand alignment and the potential for the solder to shift during
the remaining assembly. While preserving device alignment, the silicon die can then be
dropped within the graphite locator. Should there be some misalignment from the initial
drop, the edge of the die can be gently pushed using narrow tweezers, or the backside
of the package be tapped to allow the die to fall into place. The graphite lid can then
be dropped on top of the device, which only contacts the perimeter of the silicon die to
allow force to be added on top. This assembly process is visually represented in Figure
B.31, along with the addition of the weight shown in Figure B.32. The weight simply
consists of a bolt with a number of nuts attached to the end for increased mass. Higher
mass weights than the implementation shown here are typically recommended for die
attachment procedures; however, the current procedure has been reliably demonstrated.
Figure B.33: Vacuum furnace used for die attachment and getterless vacuum sealing: inside of vacuum champer with one package setup for die attachment and the second with a bimetallic temperature gauge.

Figure B.34: Vacuum furnace used for die attachment and getterless vacuum sealing: view of device and gauge through furnace window.

Figure B.35: Vacuum furnace used for die attachment and getterless vacuum sealing: applied weights to reduce vibrations from vacuum pump.
With the current setup and supplies, a batch of four devices can be loaded into the vacuum furnace at the same time. This is accomplished by loading two package assemblies on each graphite beam, as shown in Figures B.30 and B.33. Please note that due to the height of the weights, the packages must be placed at the center of the hot plate for the additional clearance that the window provides, as shown in Figure B.34.

Even though the hotplate is temperature controlled, the temperature at the location of the solder is different, due to the additional thermal load of the graphite support beam and weight on top. To help get a better idea of the temperature of the package, a bimetallic temperature gauge was placed on top of an empty package to visually monitor the temperature over the course of the temperature profile. The disadvantages of this method are the typically poor accuracy of these types of gauges, as well as the lack of the thermal load of the weight. Nevertheless, a visual indicator had to be used due to lack of electrical connection into the chamber. This method is shown in Figures B.33 through B.35, and was used to generated the measured temperature for the die attachment and vacuum sealing profiles shown in Figures 6.2 and 6.15, respectively.

As shown in Figure 6.2, the die attachment temperature profile can then be run, which heats the hotplate to 450°C for 20 min before gently cooling back to room temperature. At the start of this program, look through the window and observe the weights on top of each die. Occasionally the weights will vibrate due to the activation of the vacuum pump. Should this occur, attempt to gently add mass to the top of the chamber, as shown in Figure B.35, until there is no visible motion. When the weights vibrate, a uniform force is not generated across the die. This generally causes the die attachment to fail, ruining both the device and package.

Once the profile is complete, unload the packages. The graphite frames can easily be removes by holding the packages upside down and gently tapping on the backside over a Petri dish. An example of a successfully packaged Phase 1 QMG with front-side metallization is shown in Figure B.36.

In order to prevent damage to the vacuum furnace, please keep in mind these three points: 1) Always activate the vacuum after starting a program, 2) Always turn off the
vacuum and wait for it to purge after completing a program, and before loading a new program to the controller, and 3) Do not allow water into the vacuum chamber, which can be difficult due to condensation if the chiller is left on for a long period of time. Keeping this points in mind can prevent a significant number of damaging events from occurring.

B.2.12 Wire Bonding

The etched silicon electrodes must now be interfaced with the traces of the ceramic package. This is accomplished by wire bonding, or welding short pieces of conductive wire between the electrodes of the device and conductive bond pads of the package. This wire commonly consists of aluminum, copper, silver, or gold, and uses a combination of heat, pressure, and ultrasonic energy to make the connection. The package is designed to carry this conductive path to the outside of the package, where it can be interfaced with a specifically designed socket for easy attachment and removal. Even after the package is sealed to protect the device from the outside environment, this conductive path remains intact for electrically interfacing the device.

A number of bonding methods also exist, which dictate the final geometry of the bond, such as wedge and ball bonding. In this case, cold aluminum wedge bonding is used through the use of a Model 747677E wire bonder from WestBond, as shown in Figure B.37. Properties of the specific wire used are provided in Table B.6, along with the wire bonding settings in Table B.7.

While it is typically not recommended to bond directly onto silicon without the aid of pad metallization, no adverse effects have yet to be observed. It is possible that such influences do not appear until large scale yields are examined, with the possible addition of thermal cycling. For small-scale sensor prototyping, however, directly bonding to silicon is completely acceptable, even with the addition of thermal cycling.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item No.:</td>
<td>24909</td>
<td></td>
</tr>
<tr>
<td>Material:</td>
<td>1% Si/Al Wire</td>
<td></td>
</tr>
<tr>
<td>Diameter:</td>
<td>0.001 inch</td>
<td>25 µm</td>
</tr>
<tr>
<td>Elongation:</td>
<td>1-4%</td>
<td></td>
</tr>
<tr>
<td>Tensile Strength:</td>
<td>14-16 grams</td>
<td></td>
</tr>
<tr>
<td>Spool No.:</td>
<td>#11</td>
<td></td>
</tr>
<tr>
<td>Ft/Spool:</td>
<td>270</td>
<td></td>
</tr>
</tbody>
</table>

Table B.6: Aluminum wire properties for wirebonding.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General Settings</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loop Height:</td>
<td>50 mils</td>
<td>1270 µm</td>
</tr>
<tr>
<td>Drop Before Clamp:</td>
<td>10 mils</td>
<td>254 µm</td>
</tr>
<tr>
<td>Wire Pull:</td>
<td>34 motor steps</td>
<td></td>
</tr>
<tr>
<td>Wire Tail:</td>
<td>30 motor steps</td>
<td></td>
</tr>
<tr>
<td>Dual Force:</td>
<td>Off</td>
<td></td>
</tr>
<tr>
<td>Beep upon Contact:</td>
<td>On</td>
<td></td>
</tr>
<tr>
<td>Bond Counter:</td>
<td>On</td>
<td></td>
</tr>
<tr>
<td>Bond Counter Limit:</td>
<td>25000</td>
<td></td>
</tr>
<tr>
<td>U/S Power During Feed:</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Self-Thread:</td>
<td>Off</td>
<td></td>
</tr>
</tbody>
</table>

| **Modification for Deep Packages** |       |       |
| Drop Before Clamp: | 90 mils | 2286 µm |

| **Bonding to Gold** |       |       |
| Power:              | 245    |       |
| Time:               | 25 ms  |       |

| **Bonding to Silicon** |       |       |
| Power:                | 230    |       |
| Time:                 | 110 ms |       |

Table B.7: Wirebonder settings for bonding to gold and silicon bond pads of standard and deep packages.
When manually wire bonding, even with a semi-automatic piece of equipment as shown in Figure B.37, there are a few important things to note. First, this is a specialized skill that needs to be developed, especially when bonding open devices. Before attempting usable devices, it is best to gain practice using dummy devices with known lithography errors. Gaining a sense for how the wire behaves when drawn from the spool is necessary for preventing electrical shorts. Even after a bond is made, it is sometimes necessary to manually go into the device with a probe and, while observing under the microscope, manipulate the wires to prevent contact with the device or other wires.

Second, the optimal wire bonding settings change over time. While a majority of the settings of Table B.7 remain consistent for the specific package being used, the power and time of each bond can drift. This drift is dependant upon the cleanliness of the material being bonded, as well as the age of the bond tool being used. The bond tool is the component of the machine that places the wire in contact with the material being bonded, and resembles a needle-like structure. The bond tool is typically only useful for about 25,000 bonds in commercial settings, but this can quickly drop to 10,000 when used in academic prototyping. As the number of bonds increase, aluminum builds on the end of the tool, making future bonds more difficult. It is possible to intermittently remove this buildup using an aluminum etchant; however, if costs are not a high concern, the tool can simply be replaced. The rate of this buildup increases when higher bond powers are used, yet higher bond powers become necessary as the buildup increases. For this reason, it is best for the longevity of the tool to always use the minimum amount of power necessary for creating a strong bond. A bond power that is too high can also cause the wire to shatter during the bond, straying aluminum particles across the surface of the device. Should this occur, it may be possible to remove these particles by tapping the device upside-down, or manually removing with a probe, but this circumstance is far from ideal. When initially creating these settings, it is best to start with the recommendations in Table B.7, then slowly increasing them until an adequate bond strength is achieved. This strength can be tested by completing a dummy bond on the frame of the sensor, then attempted to manually ripping off the bond with tweezers. If the wire breaks before the interface between the wire and pad does, a suitably strong bond has been achieved.

Next, when making bonds, always allow the wire to travel from front to back, and preferably from the package to device. Bonding from front to back is critical for wedge bonding due to how the wire is fed through the bond tool. Significant deviations from this path will cause the wire to separate from the clamp behind the tool, preventing the completion of the bond and forcing the user to rethread the wire. Rethreading the wire can be a challenging task for an unexperienced user, and is a process that can be difficult to master. Bonding from the package to the device is not as critical, depending on the package dimensions. The biggest concern is that the back of the wedge travels outward at a 45° angle. When the package bond pad is the second pad bonded, and the wire is traveling from front to back, this may cause the wedge to make contact with the frame around the package before the wire bond can be completed. Sometimes this is avoidable depending on the package dimensions and the location on the pad that is being bonded. In these instances, bonding from device to package is acceptable. It is also possible to replace the 45° wedge assembly with a deep access assembly, in which case this concern
is also alleviated. When a lid is to be attached to package at a later date, which is true in most circumstances, it is also important to keep the height of the wire bonds low. This can also be a significant challenge depending on the bond pad locations on the die, as well as keeping them high enough to avoid unintentional shorting.

Finally, once the device is completely wire bonded, visually examine the device from the side under the bright microscope light to check for any bonds that may be too high and would contact the lid. Any wire bonds that can be seen must be manually lower with a probe under the microscope until they can no longer be seen from the side. It is also advised to remove the device from the chuck and tap it upside-down on a hard, clean surface. Not only does this help remove any particles that may have lodged on the surface of the device, but also insure that all the wire bonds are low enough to prevent contact with a lid. Once complete, examine the device and each of the wire bonds once again under the microscope for any shorts. Repeat as necessary. If large, stubborn particles are observed, it is also possible to feed some additional wire from the end of the wire bonder and use it to gently manipulate and lift these particles from the surface of the device. This technique can also be used to probe the structure to check for mechanical compliance.

Once complete, a label should also be etched onto package for easy sensor identification. In the past, a four character labeling system has been used: the first character signifies the batch of wafers the device was pulled from, the second character is the specific wafer within the batch, and the third and fourth characters represent the $x$ and $y$ coordinates of the specific wafer. An example of this would be “72GE” for batch 7, wafer 2, and die GE. Up to this point, only the batch and wafer had to be tracked for the devices in question. The die code is etched onto each specific device and can be read under the microscope. For DIP packages, this code can be etched onto the front of the package. For LCC packages, the code must be etched onto the back. This etching can easily be accomplished by supporting the device with one hand, while etching the code with the other using a pair of stiff tweezers.

The device can then be electrically swept using front-end electronics for characterization of the native frequency split, and to insure that both resonant axes are working. If shorting appears in the form of current limits being reached or input overloads on the lock-in amplifier, there is likely either a short due to a lithography error, particle in a critical gap, or wire bonding short. If the source of the error can be identified, some of these errors can be fixed. A breadboard implementation of the front-end electronics can be a valuable resource for identifying these issues by allowing the user to selectively rule out electrodes, one by one. For instructions on a potential implementation of this setup, please see Appendix C. Modifications to this setup can be made with the aid of Appendix A for sensor designs that significantly deviate from that of the QMG.

**B.2.13 Vacuum Sealing**

With initial characterization complete, the sensor can now be prepared for vacuum sealing. Prior to sealing, it can be beneficial to know how the sensor reacts to a reduced pressure environment. This can be used to help determine the final vacuum level within the packaged cavity after sealing occurs, and can be discovered by actuating the device
within a vacuum chamber before sealing. If low Q-factors are expected, frequency sweeps can be used to determine the Q-factor versus pressure relationship. If a high Q-factor design is tested, ring-down measurement might be necessary instead, where the device is given an impulse and the rate of amplitude to decay is measured and correlated to Q-factor. Of course, this is not a necessary step, but only if the final vacuum level within the package is to be confirmed, or if devices are being separated by their total Q-factor potential.

There are two primary methods of vacuum sealing a device: Sealing with, or without getter material. Sealing without getter material can be done in a simple vacuum furnace, such as the UniTemp furnace used for die attachment. In this case, the device must simply be placed back into the furnace on top of the graphite beams. An 80% gold and 20% tin solder frame is then placed around the device on the exposed surface of the package. For DIP packages, this solder frame has outer dimensions of 0.550 × 0.550 in, with a 0.490 × 0.490 in opening in the middle, and thickness of 0.002 in. A larger frame is required for the LCC packages due to the larger opening. A flat lid is then placed on top consisting of a gold-plated Kovar material, with careful placement to ensure that the alignment of both the frame and lid are centered. If a clear view port is desired, lids can also be fabricated using D263 glass wafers, which have a thermal expansion coefficient that matches that of the ceramic package: 7.1 ppm/K. These wafers must then be deposited with metal traces and diced to form visually transparent package lids. Weights must then be placed on the lids and a vacuum sealing temperature profile run on the controller. This profile is shown in Figure 6.15, which includes a 20 min sensor bake out at 280°C, before ramping up to 450°C for another 20 min to complete the solder reflow. Longer bake out times may be required for devices that have been stored at ambient conditions for long periods of time since the previous die attachment bake out, though the temperature should not be increased to prevent premature solder reflow. It should also be noted that maximizing the ramp rate between bake out and sealing is vital to reducing the final pressure level in the cavity. Using this technique, the final pressure level in the cavity was improved by nearly an order of magnitude, from 1 Torr to 0.3 Torr, as shown in the inset of Figure 6.15. For QMGs, these pressure levels produce Q-factors of approximately 1000. For lower pressures, getter sealing can be used.
Figure B.39: Inspex X90 X-ray System: Expanded view.

Figure B.40: Inspex X90 X-ray System: View through window.
Getter material helps to improve vacuum by continuing to absorb ambient gasses within the enclosed cavity after sealing. There are many different materials that can be used as getters, but all require a high temperature bake out prior to sealing. This bake out purges the material of these contaminants, and if sealed in an enclosed cavity, will begin to absorb molecules within it once again over time. By absorbing the remaining vapors within the cavity, the encapsulated pressure is reduced over time.

For sealing packages with an activated getter, a specialized vacuum chamber is required that consists of two heating plates and a mechanical stage. In this way, the package and lid can be baked out at different temperatures, then brought into contact afterwards to seal the device. The Model 3150 high vacuum furnace from SST International, as shown in Figure B.38, is capable of accomplishing this task and is a useful tool for getter sealing.

Independent of the sealing method, process development is difficult without the use of imaging technology to assess the presence of voids in the solder connections. This is likewise true for die attachment. For this reason, X-ray imaging can greatly aid in process development, such as through the use of an Inspex X90 X-ray system, as shown in Figures B.39 and B.40. With the images produced by this machine, the presence of voids can be reduced by either modifying the bake out time, or the sealing temperature or time, depending on the cause. Examples of these images for die attachment can be seen in Figure 6.6, with a comparison between die attachment with and without voids seen in Figure 6.1.

B.2.14 Challenges

The silicon-on-insulator process described above has been shown to be capable of producing high performance devices. That is not to say, however, that there is not room for improvement. A number of challenges continue to exist which can be summarized into three major categories: 1) Fabrication imperfections, 2) Packaging stress, and 3) Yield.

While silicon-on-insulator fabrication is desirable for a number of reasons, one of the main challenges of the process is lateral silicon etching at the buried oxide interface during DRIE etching, otherwise known as ‘footing’. Due to the high variability of etch rate as a function of feature geometry, it is nearly impossible to design the structure so that it reaches the same depth uniformly across any given device. An example of this can be seen in Figure B.16 (right), where the vertical trenches etched faster than the horizontal ones, despite an identical structural design. Designs can be modified to take this disparity in etch rate into account, such as designing critical elements, like springs, with minimal gap dimensions; however, this can only minimize the issue, but not alleviate it.

Similar asymmetry can also be observed on the wafer level. One of a primary benefits of MEMS devices is the capability of batch fabrication. Unfortunately, this is also one of the primary sources of variability between device performance. Because many devices are fabricated on a single silicon wafer, any asymmetry in the processing across the wafer results in sensor variability. As the diameter of the wafer increases to accommodate additional devices per batch, so does the potential for asymmetry across the batch. For large scale commercial processes, this may not be an issue; simply calibrate the system once and produce sensors indefinitely. However, in small scale academic settings, this
process development for each new sensor design can be challenging, resulting in high degrees of asymmetry across the wafer. For large devices, this asymmetry can even show up across a single structure. Highly symmetric structures, such as tuning forks, are challenged by this result, leading to performance deterioration.

Even though large devices are beneficial to performance by reducing thermomechanical noise, their large size also makes them more vulnerable to a wide range of yield issues. Lithography errors and particulate contamination generally occur as a function of sensor area, and it only requires one error to potentially destroy a device. For this reason, device yield is proportional to the total area of the device, but total cost is exponential.

Yield issues are not only a challenge for fabrication, but also packaging. Not only is reliably encapsulating sensors within a high vacuum a challenging task, but if the history of each package to be sealed is not maintained, it can be nearly an impossible one. As previously mentioned, bake out time must be long enough to purge contaminate from the metal, but short enough not to significantly cause metal diffusion so that a high quality gold remains on the surface. If the thermal age of the package is not known, or if there is a high variability between samples, it can be impossible to determine the optimal bake out time.

Finally, asymmetric die attachment stress can have a large impact on the innate frequency mismatch of a symmetric structure, as well as the thermal coefficient of frequency. If it is severe enough, it can even deform the suspended features of the device, as shown in Figure B.41. In order to reduce this effect, the die attachment area can be reduced; however, there is a fundamental tradeoff when high temperature solder is used. A small attachment area leads to reduced stress, but weaker attachment, while a large attachment area leads to stronger attachment, but higher imparted stress. The optimal balance between these two factors resides within the intended use of the device.

One small additional note is that the photoresist AZ4620 was used extensively in this work, including sensor definition. This was an artifact from a previous process where this
photoresist was used as an etching mask during device definition for thicknesses up to 50 µm. For this reason, a thick photoresist mask was desirable to prolong its longevity during device etching. Since then, a silicon dioxide hard mask as been incorporated, therefore the photoresist mask is only needed to protect against a few microns of etching. For this reason, a substantially thinner photoresist mask can be used. A thinner mask would allow smaller features to be defined, as well as reduce asymmetry across the wafer due to disproportionate photoresist development.

B.3 Silicon-on-Glass Fabrication Process

One of the fundamental challenges of silicon-on-insulator fabrication is an asymmetric etch rate across each device, resulting in lateral silicon etching as the buried oxide interface, otherwise known as 'footing'. Even though there are certain design choices that can be made to minimize this effect, it can not be eliminated without process modification. In this section, a bulk silicon-on-glass fabrication process is presented and discussed in an attempt to reduce this influence, while still preserving performance by maintaining a thick device layer. Experimental results are shown, which demonstrate the main challenges of the proposed fabrication.

B.3.1 Conventional Silicon-on-Glass Fabrication

As opposed to the silicon substrate of conventional silicon-on-insulator fabrication, silicon-on-glass fabrication processes utilize a thick glass substrate to suspend a silicon device layer. This involves attaching separate silicon and glass wafers, typically through anodic bonding. For future release of the suspended structures, prior to bonding, either the silicon [216] or the glass wafer [217] are selectively etched with cavities and encapsulated through the bonding process. This allows the silicon device layer to be etched after bonding, releasing the suspended structures during the etching process.

Using this method, once the device layer has been etched through, the footing effects are minimized by allowing further ion bombardment to disperse within the encapsulated cavity. If desired, an alternative method is to also embed conductive metal traces at these openings, which not only prevents footing from occurring by allowing a conductive path to the rest of the wafer, but also prevents the release of the device during etching. The encapsulated metal can then be removed later to release the structure [217].

However, these benefits would be present independent of the substrate material. Glass is specifically desirable as a substrate for a number of reasons. First, anodic bonding is known to produce a strong, high quality bond that is robust to defects. This is a useful property, especially considering the potentially small anchor areas of most devices. Second, glass is an excellent electrical insulator, which helps reduce parasitic capacitance that could otherwise overwhelm the motional signal from the device. Finally, with glass, the process can be modified to incorporate through wafer vias. This can be accomplished by reflowing the glass wafer over conductive silicon posts and lapping both sides [218]. The wafer can then be etched or bonded in a similar manner as before.
While each of these methods have been successfully demonstrated, they still require multiple processing steps and multi-mask alignment to create a simple resonating structure. In an effort to streamline the process while still alleviating asymmetry due to footing, a bulk dissolved wafer process is explored.

**B.3.2 Experimental Process Overview**

The proposed process is based on the dissolving of undoped silicon using Ethylenediamine pyrocatechol (EDP). EDP is a liquid which dissolves silicon; however, highly doped silicon can behave as an etch stop. With high doping concentrations on the order of $10^{20}$ atoms/cm$^3$, the etch rate of silicon is known to reduce by up to 50 times in EDP [219]. In the past, this technique has been used to create resonant structures by first doping the surface of a silicon wafer through diffusion, following by device etching, bonding to an etched glass substrate, and dissolving the undoped silicon using EDP [220]. However, boron diffusion is only a surface doping technique with a depth limit of approximately 25 µm [194]. For this reason, it can not be used for the creation of thick structures. It also creates an asymmetric doping profile across the thickness of the wafer, with a higher doping concentration achieved at the surface [221], potentially with magnitudes of nearly $10^{21}$ atoms/cm$^3$.

In comparison, bulk doping techniques rely on embedding the doping material in the silicon during the initial creation of the silicon boule, allowing greater doping uniformity. However, there is a reduction in the maximum doping level that can be achieved. When bulk doping is used, wafer manufacturers typically quote the doping level as a resistivity measurement in $\Omega \cdot cm$. This value can be converted to a doping concentration using Equation B.1 [222]. A plot of this relationship is provided in Figure B.42.

![Figure B.42: Boron dopant concentration in silicon verses resistivity, plotted according to Equation B.1.](image)
Figure B.43: Process flow for the experimental silicon-on-glass (SOG) fabrication process.

\[
N = \frac{1.330 \times 10^{16}}{\rho} + \frac{1.082 \times 10^{17}}{\rho[1 + (54.56\rho)^{1.105}]}
\]  

where \(N\) is the dopant density in \(\text{atoms/cm}^3\), and \(\rho\) is resistivity in \(\Omega \cdot \text{cm}\).

The lowest bulk wafer resistivity that could be found commercially was quoted as \(0.001 - 0.003 \ \Omega - \text{cm}\), as shown in Table B.1. Using Equation B.1, this translates to a doping concentration of \(3.62 \times 10^{19}\) to \(1.17 \times 10^{20}\), just barely on the verge of stopping the EDP etch. Should this doping concentration suitably reduce the etch rate, such wafers could be used to create thick, suspended structures out of single-crystal silicon, using a single mask, “footing” free, glass substrate fabrication process.

The outline of this process is shown in Figure B.43. The process begins with a silicon-on-silicon wafer with different doping concentrations: the top, thinner device layer is heavily doped for a high resistivity, while the bottom substrate will be minimally doped for future dissolving. These wafers have properties identical to the SOI wafers detailed in Table B.1, only without the buried layer of silicon dioxide. The device dimensions are then patterned and etched into the device layer, to a variable depth within the silicon substrate. Due to the lack of a buried silicon dioxide layer between the device layer and substrate, once the device layer has been etched through, the etch will continue into the substrate and not redirect the ions to produce lateral etching. There is a high tolerance in the depth of this etch, with the only concern for over etching being that of reducing the structural stability of the wafer.
Once all the features of the device layer has been etched through using a timed DRIE process, the silicon dioxide hard mask can be stripped using hydrofluoric acid. The etched features are then encapsulated, through the use of anodically bonding the wafer with an unmodified glass wafer. By soaking the wafer stack in EDP, the undoped silicon substrate is dissolved away, revealing the encapsulated features. The resonance structures can then be released using hydrofluoric acid to isotropically etch the exposed glass substrate between the etched silicon device features, much like conventional silicon-on-insulator fabrication.

Compared to traditional silicon-on-insulator fabrication, the proposed process would have a number of benefits. These primarily include the elimination of lateral etching once the buried silicon dioxide layer is reached, as well as reduced parasitic capacitances due to the insulating properties of the glass substrate. These benefits are also obtained with minimal increase in complexity, and retaining a single mask lithography process.

Process steps one through five, as denoted in Figure B.43, were successfully demonstrated. The result of this fabrication can be seen in Figure B.44, which shows an optical photograph of the resulting silicon-on-glass wafer, after the dissolving of the undoped silicon substrate. This result shows that bulk silicon doping can be used as an etch stop for EDP. However, after the wafer was diced and attempted to be released in hydrofluoric acid, most of the silicon structures immediately debonded from the glass. This occurred not only due to exposure to the acid, but simply acetone as well.

Through additional investigation, a number of interesting effects were revealed. Initial inspection revealed that in some areas across the wafer, the surface of the silicon exposed to the EDP had a number of concentrated circular features. Some of these features can be seen in the photograph of Figure B.45 (left). Optical profilometry revealed there features...
Figure B.45: Photograph of individual dies on the wafer shown in Figure B.44. Photographs taken from the top (left) and through the glass from the bottom (right). [133]

Figure B.46: Microscope photographs of the dies shown in Figure B.45. Photographs taken from the top, showing a debonded anchor sitting upside-down on top of the proof mass (left), as well as from the bottom, showing anchors that remained attached (right). In both images, significant undercutting of anchors can be seen.
to not be simply discolorations due to contaminants, but rather plateaus of silicon. It is believed that in some area, when the encapsulated features were finally opened to the EDP bath, the air collected on the surface of the wafer, preventing further EDP exposure. This resulted in stopping the etch in these areas, producing the circular silicon mounds. In order to prevent such behavior in any future processes, it is recommended to ultrasonically agitate the wafers during the EDP emersion.

A second important note can be observed in both the photograph of Figure B.45 (right), as well as the microscope images of Figure B.46. During the EDP etching step, a significant level of undercutting had occurred to the silicon anchors. The amount of this undercutting can be correlated to the width of the surrounding trenches, and appears to be the primary cause of anchor debonding. This undercutting is very similar to the footing effects caused by DRIE etching of conventional SOI wafers, and only occurs in the first few microns of silicon at the glass interface.

The reason for this effect may be due to anodic bonding. While anodic bonding is known to produce a strong wafer bond between silicon and glass, part of the reason for this is due to the transfer of ions in the strong electric field. One hypothesis is that the doping level of the silicon was depleted during this bond, but only at the glass interface. When exposed to EDP, this depletion zone was unable to substantially resist the etch, resulting in undercutting and a weakened anchor attachment. This technique is successful when surface doping is used due to the higher doping concentration that is generated, especially at the bonding interface.

B.3.3 Challenges

Through experimental investigation, a number of challenges have been identified for the given bulk silicon-on-glass fabrication process. This includes: 1) A maximum bulk doping level which borders on behaving as an etch stop for EDP, 2) The undercutting of the silicon anchors when the silicon features are exposed to EDP, possibly due to a doping depletion zone from anodic bonding, and 3) Bubble formation which inhibits the symmetry of the EDP etch.

To eliminate the formation of bubbles as the device layer becomes exposed, two techniques could easily be employed: 1) Complete the wafer bonding step in vacuum, or 2) Ultrasonically agitate the wafer towards the end of the EDP etch. This challenge is not believed to be of significant difficulty. Of perhaps greater concern is the maximum level of bulk doping and the undercutting of anchors during the EDP etch.

One method of possibly solving these issues is by combining both bulk and surface doping techniques. By initially starting with a silicon-on-silicon wafer with bulk doped device layer, then further enhancing the doping concentration of the top surface of the device layer, it may be possible to prevent the undercutting of the anchors as they become exposed. This would maintain the strong ionic bond between the silicon and glass, as has been confirmed by past published dissolved wafer processes. The final step would be to release the devices with exposure to vapor-phase hydrofluoric acid, which has yet to be untested.
B.4 Conclusion

Fabrication tolerances are of critical importance for high performance vibratory sensors. Conventional silicon-on-insulator (SOI) fabrication is an excellent approach if rapid prototyping is the goal, and is ideal for attaining proof-of-concept of initial sensor designs. However, high-performance sensors require a high degree of uniformity that is difficult, if not impossible to reliably obtain with this process.

There are a wide range of alternative fabrication approaches that have been attempted within literature, each with varying results. However, for high-performance vibratory sensors, fabrication uniformity is a major goal. With respect to conventional SOI fabrication, the challenges include: 1) Elimination of the footing effect once the buried oxide layer is reached, 2) The creation of uniform, straight, high aspect ratio trenches, and 3) Fabrication uniformity across the entire wafer for reliable uniformity within each batch of devices.

The first challenge requires a change in material. Within the literature, attempts have been made to encapsulate empty space, or an electrically conductive layer of metal beneath the opened trenches. The concern with the first approach is the accumulation of heat in the device layer during etching, which changes the etching parameters across the wafer. Attempts have been made to include a sacrificial material in order to prevent this from occurring, but the effectiveness of this technique, along with the second approach, is still uncertain.

The second and third challenges require any number of the following methods: 1) DRIE process optimization, 2) Etch post-processing to reduce gaps or smooth sidewalls, or 3) New fabrication technology. There have been many published attempts of modifying DRIE recipes to achieve higher aspect ratios or gap uniformity. Post-processing has also been used as a means to reduce gap width through epitaxial silicon growth, and hydrogen annealing may also be used to smooth sidewall roughness, at the cost of rounding device corners [223]. There is also the possibility that there are alternative fabrication techniques to DRIE etching, such as micro-scale glass blowing, that can improve structural symmetry. However, as with all new processes, there are numerous challenges that must be overcome when converting to new forms of fabrication; for the given example, this can be thought of as a high risk / high reward tradeoff.
Appendix C

Protocols

The following guides are written with the intent that they can be printed and used as stand-alone protocols for some of the more basic procedures concerning MEMS fabrication and testing. The author makes no assumptions concerning the experience of the audience, with the intention they can be utilized as quick-start guides, reducing the learning curve for novices who desire to quickly fabricate and characterize new inertial sensor designs. The process is separated into four guides:

1. Silicon-on-Insulator Fabrication
2. Packaging of Silicon Dies
3. Breadboard Front-End Electronics
4. Open-loop Gyroscope Operation using a Zurich HF2LI Lock-in Amplifier

The first guide details the wafer-level processes required to create individual, released silicon structures from design conception. The second guide shows how these structured can be packaged for electrical interfacing. The third guide demonstrates how to build front-end electronics for interfacing the MEMS. The fourth guide explains how to force a resonator with a variable frequency into a fixed amplitude of motion, specifically for open-loop gyroscopes.

These guides only cover enough material for successful implementation. For additional theory and details on how to modify these processes for specific purposes, please see Appendices A and B. When purchasing the components required by each guide, please consult the Vendor list of Appendix D.
C.1 Silicon-on-Insulator Fabrication

This is a step-by-step guide to fabricating electrostatic MEMS using the silicon-on-insulator fabrication method. This technology produces suspended silicon structures, with the addition of backside metallization for die attachment. Many pieces of equipment are needed for this process and each requires specialized training to operate. This guide does not serve as an adequate guide for operation of these individual pieces of equipment and only supplies the critical details necessary for repeating the existing process.

The following materials are required: 4-inch silicon-on-insulator wafers, individual wafer holders (one per wafer being processed), cleanroom toolbox, wafer tweezers (2 pair), die tweezers, liquid containers (three, with at least one made of glass), hotplate, targets for metal deposition (chrome and gold), photoresist (AZ4620), thermometer, Petri dishes (one per batch), and Texwipes (many). The equipment used in the process is spread out over three cleanroom facilities, not including vendors: the INRF at UCI, the Nanolab at UCLA, and the MicroSystems Lab at UCI. The user needs training on the following pieces of equipment: Wet Benches (INRF), Positive Photoresist Spinners (INRF), Karl Suss MA6 Mask Aligner (INRF), Temescal CV-8 E-Beam Evaporator (INRF), Dehydration Oven (INRF), YES HMDS Prime Oven (INRF), Matrix Asher (Nanolab), STS AOE Advanced Oxide Etcher (Nanolab), Nanospec (Nanolab), Plasma-Therm FDRIE (Nanolab), Dicing Saw (Vendor), Scanning Electron Microscope (UCI), Vapor-phase HF Etcher (MicroSystems Lab), and Idonus IR Microscope (MicroSystems Lab). The user should also be trained in software for finite element analysis (COMSOL, Ansys, etc.), as well as L-Edit for photolithography mask creation. For additional details or specifications, please see Appendix B.

C.1.1 Sensor Design and Mask Fabrication

1. Design sensor geometry within L-Edit, or other mask creation software.

2. Export sensor design to GDSII. Before exporting, each of the following can improve the conversion process: Fill any etch holes, flattening the structure, and merge the structure geometry. Please note not to save the file after completing these actions.

3. Import the GDSII into COMSOL, or other finite element analysis software. Converting the file from GDSII to DXF using software such as LinkCAD before importing the file can improve the process.

4. Model the structure for modal frequencies and other relevant features, such as thermomechanical damping.

5. Repeat the process making alterations until desired qualities are met. Small modifications can be made in COMSOL; however, larger modifications should be made in L-Edit and the file converted once again.

6. Finalize the sensor design with L-Edit and form an array across the area of a four inch diameter silicon wafer. If multiple masks will be used, make sure to include
alignment marks approximately one inch in from the left and right sides of the wafer. Including a unique die code on the frame of each device is recommended for preserving the location it was pulled from the wafer after packaging.

7. Convert the device array to GDSII and submit the file for mask fabrication. Photo Sciences Incorporated is typically used for this fabrication, from which chrome masks are created on a soda lime substrate.

Note:

C.1.2 Backside Metal Deposition and Patterning

1. Acquire silicon-on-insulator wafers. Open container in cleanroom fume hood and place a batch of wafers into individual wafer carriers (typically five for metal deposition). If front-side gold is desired, the wafers should have no front-side hard mask initially.

2. Strip native oxide on backside of wafers. Prepare a 2% HF bath and soak each wafer for 15 seconds. Remove wafer, rinse with DI water, submerge in DI water, and blow dry.

3. Deposit chrome and gold on backside of wafers. Immediately after stripping the native oxide, load the wafers into the E-Beam evaporator and begin pumping down the chamber. It is recommended to begin this process shortly before the INRF closes so that it can pump down overnight and deposition can be done the following morning. Proceed to deposit 50 nm of chrome and 500 nm of gold with the following rates:

   (a) Cr: From 0 to 10 nm, 0.01 nm/sec.
   (b) Cr: From 10 to 50 nm, 0.03 nm/sec.
   (c) Au: From 0 to 10 nm, 0.01 nm/sec.
   (d) Au: From 10 to 50 nm, 0.05 nm/sec.
   (e) Au: From 50 to 450 nm, 0.1 nm/sec.
   (f) Au: From 450 to 500 nm, 0.05 nm/sec.

4. Clean wafers. Complete a solvent rinse using acetone and isopropyl alcohol (IPA). Hold the wafer above a liquid container and flood both surfaces with acetone, following by IPA, without the wafer drying. Keep tweezers located near the bottom to prevent contamination. Immediately rinse with DI water and dry with nitrogen.

5. Dehydrate wafers. Place in the dehydration oven for 1 hour @ 120 °C. Afterwards, blow dry with nitrogen to remove any particles.
6. Complete backside lithography. Spin photoresist AZ460 on the wafer in two steps: 10 sec @ 500 rpm and 2 acceleration, then 40 sec @ 5000 rpm and 6 acceleration. Soft bake for 20 min @ 90°C. Expose using desired mask with a soft contact and 25 µm spacing for 36 sec. Develop in a 3.5 : 1 DI water to developer solution. Do not agitation the solution.

7. Etch gold on backside of wafers. Prepare two baths: one of DI water and a second of gold etchant. Quickly submerge wafer with pattern facing upwards in etchant. Leave wafer in the liquid for 2 minutes, agitating at least every 30 seconds. Briefly remove wafer from the liquid to observe if gold has been complete removed. If not, continue etching. Total etch time varies over time. When the etch is complete, rinse wafer with DI water in sink and submerge in the DI water bath. Observe wafer again to insure all gold has been removed.

8. Etch chrome on backside of wafers. The etching process is identical to that of gold etching.

9. Clean wafers. Wet cleaning using RCA-1 or Piranha is recommended.

Note: The gold and chrome etchants can be reused, so do not dispose after use. The only exception to this is if the solution is dirty or etch time has already been significantly reduced. For this reason, it is useful to store the liquids in a plastic containers with an attachable lid.

C.1.3 Front-side Metal Deposition and Etching (optional)

1. Strip native oxide on front-side of wafers. Prepare a bath of 2% HF and etch for 15 sec, as previously described.

2. Dehydrate wafers. Place in the dehydration oven for 1 hour @ 120 °C, as previously described.


4. Deposit chrome and gold on front-side of wafer. Chrome and gold thicknesses are 50 nm and 500 nm, respectively, and deposit as previously described.

5. Soak in solvent to remove photoresist and lift off the unnecessary metallization. Can use the photoresist developer or acetone.

6. Clean wafers. Complete a solvent rinse, as previously described, along with an aggressive high temperature process, such as RCA-1 or Piranha.

7. Deposit silicon dioxide on the front-side of wafers using chemical vapor deposition (CVD).
Note: Always perform the backside patterning before the front-side when applying front-side metallization. There are two reasons for this: 1) Backside alignment is significantly easier than front-side alignment, and 2) When using more than two masks, if possible, always align all masks to only one to help reduce tolerance errors.

C.1.4 Front-side Patterning and Etching

1. Dehydrate wafers. Place in the dehydration oven for 1 hour @ 120°C. Afterwards blow dry with nitrogen to remove any particles.

2. Apply a layer of HMDS in the “YES” oven, with a processing time of approximately 30 min. This step promotes photoresist adhesion and is required when patterning silicon dioxide.

3. Complete front-side lithography. Spin photoresist AZ4620 on the wafer in two steps: 10 sec @ 500 rpm and 2 acceleration, then 40 sec @ 5000 rpm and 6 acceleration. Soft bake for 20 min @ 90°C. Expose using desired mask with a soft contact and 25 µm spacing for 36 sec. Develop in a 3.5 : 1 DI water to developer solution. Do not agitation the solution.

4. Etch silicon dioxide hard mask on the front-side of wafers. This is performed using the STS AOE in the UCLA Nanolab, with recipe “oxidapic”. Etching takes approximately 4 min; however, must be monitored throughout the etching process using the Nanospec.

5. Strip the photoresist with O2 plasma. This is completed in the Matrix Asher using the recipe “3 min STRIP”. Run this recipe four times for each wafer, for a total of 12 min of stripping to ensure that all of the photoresist has been removed.

6. Etch silicon device layer. This is completed in the FDRIE using the recipe “Darkfield2”. Total etch time can vary between 30 and 35 min for a 100 µm thick device layer.

7. Strip wafers with O2 plasma. Once etching is complete, clean the wafers once again using O2 plasma with the Matrix Asher. Repeat the prior process by running the recipe “3 min STRIP” four times, for a total of 12 min of stripping.

Note:

C.1.5 Wafer Dicing

1. Deposit a protective coating to the front-side of wafers. Spin AZ4620 as if performing lithography, but change the speed of the second step from 5000 rpm to 2000 rpm, and reduce the soft bake from 20 min to 2 min.

2. Dice wafers into individual dies using a dicing saw. If SEM inspection is desired, include an additional cut across one or more sacrificial dies for imaging. A dicing diagram should be provided to the vendor with detailed dimensions.
Note: Alternatively to mechanical dicing, laser dicing is also possible and is probably preferred.

C.1.6 Scanning Electron Microscopy (optional)

1. Remove the sacrificed die from the holder and clean. This step can be performed at the same time as a batch of whole dies to save time. As described in the next section, this involves an acetone soak followed by an aggressive wet cleaning with RCA-1 or Piranha.

2. Mount the die for imaging. SEM mounts are required for this step, along with conductive tape. Specific mounts are required for mounting the sample vertically to observe the diced cross-section.

3. Inspect the cross-section for etch quality. Operate the microscope to image the cross-section at any number of critical locations, especially springs.

Note:

C.1.7 Die Cleaning and Release

1. Blow dry diced wafer. Open the wafer holder inside a clean room fume hood and gently blow dry the diced wafer to remove any loose particles on the surface.

2. Prepare two baths of acetone and one bath of deionized water.

3. Remove dies from the holder. Lay down a clean Texwipe and hold the wafer upside down over it. Gently press on the back of the plastic until a die dislodges and falls to the wide.

4. Hold each individual die upside-down in the first acetone bath. Do not release inside the first acetone bath; just hold it upside-down in the liquid for 15 to 20 sec to allow the photoresist to dissolve and pull any encapsulated particles to the bottom of the bath through gravity.

5. Place each individual die right-side-up in the second acetone bath and soak for 10 minutes. This is to dissolve the bulk of any additional photoresist before further aggressive cleaning.

6. Remove dies from the second acetone bath and place into the deionized water. This is to act as a buffer before further aggressive cleaning, so acetone is not added to the next bath.

7. Prepare a bath of RCA-1 or Piranha and clean the dies for 10 to 15 min. The purpose of this step is to ensure the devices are clean prior to release and packaging. Place the dies into a fresh bath of DI water once the cleaning time has elapsed.
8. Transfer the dies from the DI water bath to a bath of IPA and soak for at least 2 min. IPA evaporates significantly faster than water, resulting in reduced drying times.

9. Transfer the dies from the IPA bath and place on Texwipe. Gently shake IPA from devices when removing from the bath to reduce the drying time. Cluster the devices together on the Texwipe, face-up, so that a cover can be applied in order to reduce the change of particles falling on them. Prop the cover up on one side to allow diffusion to occur.

10. Wait for the IPA to visibly dry from the surface of the devices, which takes approximately 10 min. Transfer each die to another part of the Texwipe and wait another 5 min. This is to ensure that the back of the dies is also dry.

11. Once the \( \approx 15 \) min drying time is complete, place each die in a Petri dish.

12. Take the devices from the cleanroom to the vapor HF machine and etch the devices at 36\( ^\circ \)C for \( \approx 45 \) min. Etch time increases as the concentration of the liquid HF reduces.

13. Once complete, check release with the IR microscope. Repeat etch until the buried oxide is ideally removed.

Note: It is critical that the dies be dry before placing into a Petri dish. Even if they are dry, the two smooth surfaces can adhere very well on their own. For this reason, be very careful when removing devices from a Petri dish and tap them from the side to ensure they are not stuck before attempting to lift them out. If devices are stuck, first remove the loose devices. Next, attempt to push the devices from the side to shear the attachment. If this fails, turn the Petri dish upside down on a clean Texwipe and push from the back. This will work, but it is not ideal due to the front-side exposure.

C.1.8 Infrared Microscope Inspection

1. Turn the computer on and login. The power button is on the lower left side of the screen. Press it and wait for the login menu to appear. Type in the password and wait for the desktop appears.

2. Turn the microscope on. There is a switch on the front panel of the microscope at the left labeled “Main”. Switch the setting from “0” to “1”. Either a red or green light will turn on next to this switch when the microscope is active, indicating the mode it is in. Wait for computer to detect the microscope; several audible chimes will sound when it does so.

3. Manipulate the microscope for either IR or visible light imaging. There are three physical settings on the microscope that must be manipulated when switching between IR and visible light: 1) A switch on the front panel of the microscope, to the right of the “Main” switch, 2) A lever on the left side of the microscope, and 3) A
slider on the right side of the microscope. For IR light, these settings must be set to either “Transmission” or “IR”. For visible light, these settings must be set to either “Reflection” or “Visible”.

4. Open the microscope imaging program. Click on the start menu, then the program “INFINITY ANALYZE”. Wait for it to open. If an error box pops up, click “OK”, close the program, make sure the microscope is turned on, then reopen the program. When the program initially starts, it automatically attempts to connect to the microscope, so an error pops up if it is unable to do so.

5. Place sample under the microscope. The previous user should have placed a Petri dish over the sample holder; remove the cover. For visible light, it does not matter where you place the die, so simply set it somewhere on the stage. For IR light, IR cannot pass through the stage, therefore various sized holes have been cut into it. Place your sample above one of these holes, making sure it is completely covering the chosen hole.

6. Manipulate the software settings, source intensity, and focus until the sample is visible.

7. Adjust sample until chosen region is visible. The stage provides some x- and y-axis control; however, due to the opaque stage, the device usually has to be manually repositioned with tweezers.

8. Save the image and copy to an external drive to permanent saving.

9. Close the software and turn off the computer. The software can simply be closed by clicking the “X” in the upper right hand corner. Turn off the computer through the Start menu.

10. Turn the microscope off. Flip the switch on the front from “1” to “0”. This is a critical step to preserve the life of the IR source.

Note:
C.2 Packaging of Silicon Dies

This is a step-by-step guide to packaging silicon dies using the Uni-Temp RSS-450 Vacuum Furnace for die attachment and the West-Bond Wedge Wire Bonder for wire bonding. Vacuum sealing without getter material is also an optional process using the vacuum furnace. The following items are required: Vacuum furnace and supporting equipment (Furnace, PC, controller, power source, vacuum pump, and chiller), wire bonder and supporting equipment (wedge tip and aluminum wire), silicon dies with backside gold for attachment, DIP or LCC packages, Au/Sn solder preforms, graphite locators, graphite bars, and weights. For additional details or specifications, please see Appendix B.

C.2.1 Turn On Equipment

1. Plug in the power cord to the chiller (located above the wire bonder).

2. Confirm that the water level in the chiller is adequate. The indicator is on the lower left of the front panel of the machine. If the water level is below “Min”, add additional water to the black inlet.

3. Turn on vacuum pump power by flipping the green switch on the back. Do not start actively pumping.

4. Turn on the power supply to the vacuum furnace controller by flipping the white switch on the front. It is the black transformer located underneath the setup, as shown in Figure C.1.

5. Turn on the vacuum furnace controller using the green switch on the front of the unit.

6. If off, turn on the computer. Username and password are on a sticker on the keyboard.

7. Run the UniTemp program on the computer. The software will make a connection to the controller.

Note: A valve controlled by the vacuum furnace controller separates the vacuum pump from the chamber. The default position of this valve is closed. Never activate the pump when this valve is closed. The pump connected to the chamber has two components: a roughing pump for atmospheric pressure and a turbo pump for low pressures. The turbo pump can be damaged when exposed to atmospheric pressure, and does not activate until the pressure in the line is below a certain level. Running the pump when the valve is closed will reduce the line pressure and activate the turbo pump. Should the valve suddenly be opened (such as when a program is run), the turbo pump would instantly be exposed to atmospheric pressure, thus resulting in damage.
C.2.2 Bake-out the Chamber (optional)

1. Clean the interior of the vacuum furnace with IPA. Using a low particulate clean room wipe, moisten with IPA, and wipe the interior of the vacuum furnace. Wipe every interior surface that can be reached, along with the graphite bars that will be used for future procedures. Wait for any residual IPA in the chamber to evaporate before continuing.

2. Close the vacuum furnace and lock the handle. When properly locked, the handle will clink with metal against metal.

3. Open the program screen. Using the UniTemp program on the computer, select “Program” at the top of the screen. A new window will open for editing controller programs.

4. Open the chamber bake-out program. Select “File” → “Open program file”. A pop-up window will open and select “1-ChamberBakeOut” and click “Open”. The saved program will appear on the screen and will be available to edit.

5. Load the program to the controller. Select “Device” → “Save program to device”. A pop-up warning will appear. Click “OK”. Wait a few seconds until the controller stops clicking.

6. Close the program screen by clicking the “X” in the upper right corner.

7. Open the monitor screen. Select “Monitor” at the top of the screen, and a new window will open for running and observing the loading program during the process.
8. Run the program. In the lower right of the monitor screen, text will display the file path to the program that was loaded onto the controller last, in this case “1-ChamberBakeOut”. Above this text are the buttons labeled “Run” and “Reset”. Click “Run” to start the program. Note that at any time during the run, you can click “Reset” to stop the program.

9. Activate the vacuum pump. Immediately after running the program, the circle next to the label “Vacuum” under the “Digital Output” on the computer screen will turn from grey to red, indicating that the valve between the pump and vacuum chamber is open. Activate the pump by pressing the soft “Power” button on the front panel of the vacuum pump.

10. Wait for the program to finish. The program takes about 2 hours to run. Being present until the chamber reaches the maximum temperature is required, as sometimes the pump deactivates when temperature rises too fast. Should this happen, press the soft “Power” button on the front panel of the vacuum pump again, until the pump activates.

11. Deactivate the vacuum pump by pressing the soft “Power” button on the front panel. Do not run another program until the pump purges the line, which takes about 5 minutes.

Note: Baking out the chamber is an optional process, but should be done routinely. If you are vacuum sealing a device, the procedure is not optional and is required.

C.2.3 Back-out Packages

1. Select the packages that will be used. If DIP packages are being used, first cut off the metal connecting the pins using tin snips, as shown in Figure C.2.

2. Open the vacuum furnace and place the packages and dies inside for bake-out. Packages should be placed on the graphite blocks, while the silicon dies are placed directly on the hotplate, propped up by the graphite locators. This setup is shown in Figure C.3 for a batch of four dies.

3. Close the vacuum furnace and lock the handle (Please see the previous section if more information is required for this and the following steps).

4. On the computer, open the program screen.

5. Open the package bake-out program: “2-PackageBakeOut”.

6. Load the program to the controller.

7. Close the program screen.

8. Open the monitor screen.

9. Run the program.
10. Activate the vacuum pump.

11. Wait for the program to finish. The program takes about 3.5 hours to run.

12. Deactivate the vacuum pump. Do not run another program until the pump purges, which takes about 5 minutes.

Note: The die attachment program should be run immediately after the package bake-out program, with minimal delay.

C.2.4 Die Attachment

1. Correctly orient the silicon dies. Remove the silicon dies from the vacuum furnace and place into a Petri dish and examine under a microscope. Observe the die code in the upper left and right corners of each die and orient so that the direction is known, even without the microscope.

2. Insert the graphite locator frames into the cavity of each package. See Figure C.4 (left).

3. Insert a small solder preform into the center of each cavity. If need be, probe with tweezers until the solder preform is as close to the center as possible using manual manipulation. See Figure C.4 (left).

4. Insert the silicon die so that the top of the die is oriented along Pin 1 of the package. See Figure C.4 (center). For DIP packages, Pin 1 is on the same side as the metal trace leading away from the cavity seal (to the left, as shown in Figure C.4). This is important for being able to preserve a standard wire bonding scheme, as well as tracking wafer-level data. If the silicon does not immediately fall into the cavity, the backside of the package can be tapped to aid in placement.
5. Insert the graphite lid. See Figure C.4 (right). Note that there is a shallow cavity milled into one side of the lid, forming a graphite rim around the parameter. This side must be placed in contact with the silicon die so that only the parameter of the die is in contact with the graphite lid. If the lid does not immediately fall into the cavity, the backside of the package can be tapped to aid in placement.

6. Place the package assembly into the vacuum furnace, onto the graphite beams. The graphite beams should be centered on the hot plate in the vacuum furnace.

7. Repeat steps 2 through 6 for each die being packaged.

8. Carefully place weights onto each graphite lid. See Figure C.5. Bolts are currently used with a number of nuts attached to each. Make sure the nuts are screwed against each other so that they are not loose on the bolt.

9. Close the vacuum furnace and latch the handle. If need be, carefully slide the graphite beams so that the bolts are within the viewing window. The height of the bolts may interfere with the lid and the window offers additional clearance. See Figures C.6 and C.7.

10. On the computer, open the program screen.

11. Open the package bake-out program: “3-DieAttach”.

12. Load the program to the controller.

13. Close the program screen.

14. Open the monitor screen.

15. Run the program.
16. Activate the vacuum pump.

17. Observe the die attachment weights through the window. If one or more are vibrating, add additional weight to the top of the vacuum furnace until the vibration stops. Without a constant pressure being applied to the die, the attachment will not be successful. See Figure C.8.

18. Wait for the program to finish. The program takes about 1 hour to run.

19. Deactivate the vacuum pump. Do not run another program until the pump purges, which takes about 5 minutes.

20. Open the vacuum furnace and remove packages. After removing weights, grip the package with tweezers and flip upside down over Petri dish. The graphite lid and frame should fall due to gravity. If they do not, grip package with fingers and tap on the back until the graphite is free.

21. Place packaged dies in a separate Petri dish. Dies should always remain covered, even in the cleanroom, to avoid unnecessary exposure to particles. Resulting packaged dies should resemble Figure C.9.

Note:

C.2.5 Wire Bonding

1. Turn on the wire bonder. The West-Bond wire bonder is shown in Figure C.10. Three switches must be flipped: the toggle on the top box on the left side of the microscope, the toggle on the second box down, and finally the toggle on the microscope that turns on the microscope light.

2. Affix the packaged device to the chuck. The chuck has multiple grooves cut into it, with a clamp on one side. A button on the base opens and closes the clamp. If wire bonding a DIP package, open the clamp, place the package onto the chuck next to the clamp so that the leads drop into the grooves, and close the clamp. If wire bonding an LCC package, an empty DIP package will be required to brace the LCC between the clamp and the DIP package.
Figure C.5: Weights placed on top of the graphite lid to apply pressure to bond.

Figure C.6: Die attachment assembly inside the vacuum furnace, with bimetallic temperature gauge for measuring package temperature.
Figure C.7: Temperature gauge and weights visible through observation window in vacuum furnace.

Figure C.8: Additional weights applied to the top of the vacuum furnace for reducing vibrations.
3. Adjust the height of the stage so that the wire bonding tip has adequate clearance to the package. This can be accomplished by adjusting the knob on the front of the wire bonding stage.

4. Label the package. Look at the silicon device through the package to determine the die code. This code, along with a code representing the wafer it came from, can be etched into the package using thick tweezers. It is important to keep detailed notes so that at any point in time, the entire history of a silicon device can be traced back to conception. This is useful information for process optimization. Here is an example of QMG labeling: “72GH”. This represents batch “7”, wafer “2”, device “GH”.

5. Choose your settings by selecting the correct buffer. The current buffer, or group of settings, is shown in the upper right of the wire bonder display. The buffer can be changed up or down using the second toggle labeled “BUFFER”. For each buffer, the main display shows the number of bonds in the buffer, the “Power” and “Time” for the first bond, and the number of bonds since the tip was last changed. Further details of the buffer can be displayed by pushing the fourth toggle down to “EDIT” repeatedly. At any point in time, the main screen can be returned by pushing the fourth toggle up to “US/ESC”. If you are a new user, speak with current users to see which buffers are occupied. Do not change the setting of another person’s buffer, but rather a new, unoccupied buffer should be assigned to you.

6. Modify the buffer, if need be. Attempt a few test bonds to see if your current settings work correctly. As bonding tips age, it is sometime necessary to continually adjust the power and/or time of one or both bonds. The power and time determine the bond strength: “Power” is the level of ultrasonic power that is delivered by the bond tip, while the “Time” is the length of time of the bond in milliseconds. If the
power and/or time are too high, aluminum particles can spray onto the device and reduce the number of bonds before the tip must be replaced. If the power and/or time are too low, the bond will be weak. To test the bond strength after completing a test bond, attempt to remove it with a thin probe. If the wire breaks before it detaches from either bond location, the bond strength is excellent. As a rule of thumb, use moderate power (245) and low time (25) to bond to gold bond pads, and low power (230) and high time (110) to bond to silicon anchors. To change any settings, simply use the third toggle labeled “UP” and “DOWN”.

7. Wire bond the device. For a given device design, a standard wire bonding layout should be established beforehand. This allows custom front-end electronics to be designed and built for a given design, rather than a single sensor, and allows the user to be confident of the wire bonding layout, even if opaque lids are used during vacuum sealing. Also note that when wire bonding, all wire bonds should be made from front to back, as straight as possible, and from package to device. This helps prevent the wire from become unthreaded. Note that each wire bond must also be low enough not to touch an lid after sealing is completed, but high enough so as to only touch the device at the indented location. This can be very difficult and takes time to master.

8. If the wire becomes unthreaded during wire bonding, rethread it. The wire can fall out of the bond tip for a number of reasons. To rethread, press the first toggle upwards towards “THREAD/NEXT”. This opens the clamp and allows the wire to be fed through. Next, grab the free wire with tweezers and pull it forward until a straight section of wire is removed. Cut off the bent wire with scissors. Grab the wire with small tweezers a few millimeters back from the tip and attempt to feed it through the back of the bonding tip. Because the hole is small and facing away from the user, this is a skill that can take a lot of time to develop. Once the wire is through the hole, feed it through until it emerges from the bottom to a length that can be gripped by tweezers. Grab the free end and pull it through so that the wire is taught. Look back on either side of the clamp. Many times the wire will not pass through the clamp, but will rather pass around one of the sides. Push the wire back towards the center of the clamp and pull the wire through again. Repeat until the wire passes through the clamp. Press the first toggle downwards towards “TORCH/PREV” to close the clamp. The excess wire can then be bonded to test location, and wire bonding can continue.

9. Once wire bonding is complete, remove the device from the chuck and place in a save, covered location. This device must now be tested electrostatically. After wire bonding, it can also be useful to turn the device upside down and either tap on the back with tweezers, or gently tap the device onto a hard, clean surface. This can help remove any dust or particles that might prevent resonance or cause electrical shorts.

10. Shut down the wire bonder. Turn off the toggle for the microscope light, the power toggle on the second box down from the top, and finally the top box.
Note: Wire bonding is a skill that takes some practice. When starting out, do not attempt to wire bond the best devices, but rather use dummy dies until you feel comfortable.

C.2.6 Vacuum Sealing (optional)

1. Electrostatically sweep the device to insure operation in air. Non-operational devices should never be vacuum sealed.

2. Follow the protocol outlined in the first section in this document entitled “Bake-out the Chamber”. This will ensure a clean environment for sealing.

3. Place the device on a graphite beam inside the vacuum furnace. The device should be centered on the hot place inside the chamber.

4. Insert solder preform. A solder preform is then placed onto the package. Center this frame as evenly as possible.

5. Insert lid. If a Kovar lid is used, the lid can simply be placed onto the package. If a glass lid is used, remove a single lid from the diced wafer and clean with IPA. Blow dry with compressed air and place onto the package. It is critical that the solder preform and lid be centered on the package as precisely as possible. Because two items must be aligned together, and overlapping, it can be a tedious process and takes practice. If it is centered correctly, the lid should completely envelop the solder frame.

6. Gently place a die attachment weight on top of the lid. Be careful not to disrupt the alignment of the lid when doing so. Additional weights and/or materials can be added as well to help keep the lid centered during the process. The largest factor
that reduces yield is vibrations from the vacuum pump causing the lid to become misaligned.

7. Close the vacuum furnace and latch the handle.
8. Carefully place weights on top of the vacuum chamber to help reduce vibrations.
9. On the computer, open the program screen.
10. Open the vacuum sealing program.
11. Load the program to the controller.
12. Close the program screen.
13. Open the monitor screen.
14. Run the program.
15. Activate the vacuum pump.
16. Wait for the program to finish.
17. Deactivate the vacuum pump. Do not run another program until the pump purges, which takes about 5 minutes.
18. Open vacuum furnace and remove packages.

Note: Vacuum sealing can be performed in the vacuum furnace, but this procedure is much better handled using the SST International High Vacuum Furnace Model 3150. This equipment is specifically designed for vacuum sealing and produces higher yield, as well as offers the ability to seal with getter material for ultra-low vacuum cavities.

C.2.7 Turn Off Equipment

1. Close the UniTemp software on the computer. Depending on which windows are open, they may require multiple clicks of the “X” in the upper right of the screen. Close each window until the desktop is displayed.
2. Turn off the vacuum furnace controller using the green switch on the front of the unit.
3. Turn off the power supply to the vacuum furnace controller by flipping the whist switch on the front.
4. Turn off the vacuum pump power by flipping the green switch on the back of the unit.
5. Unplug the power cord to the chiller (located above the wire bonder).
Note: Of all the equipment, it is most critical to unplug the chiller after use. Occasionally, the dew point of the environment is above the temperature of the chiller, which eventually causes condensation to accumulate on the vacuum furnace if left running for a long period of time. Should this occur, this water should be wiped up using clean room wipes to avoid moisture entering the vacuum furnace when opened. Water inside the furnace will evaporate when a program is run and will enter the vacuum pump, which can cause damage.

C.2.8 X-Ray Inspection

1. Turn the equipment on. First, turn the key on the front of the system. Second, flip the switch next to the key to the “Power” symbol. Third, press the “Power” symbol button on the computer next to the equipment. Wait until the login screen appears.

2. Log in to the equipment using the password. Wait for the user settings to load.

3. Run the control program. There is an icon on the desktop labeled “X-ray control”, which can be double-clicked.

4. In the control program, click “View” → “Stay On Top”. This allows the controls to always remain visible.

5. Run the viewer program. There is an icon on the desktop labeled “XVu 1-04-2010”, which can be double-clicked. A large window will take up the entire screen. Wait for the machine to complete an automatic homing process.

6. If it has been more than a week since last use, condition the tube. This is completed in the control program by choosing “System” → “Tube Conditioning, 15 min. 90KV”, then immediately pressing the large “On” button on the control program. The x-ray tube will slowly ramp to 90 kV over the course of 15 min. Once complete, the x-rays will automatically disengage. Wait for this process to complete before continuing (The red light on top should be off).

7. Load the sample into the machine. Open the door and attach the sample to the movable arm. Be carefully of how the sample is attached; the clamp is not attached well into the arm. For delicate structures, a Petri dish can be clamped and the samples suspended on the dish. Close the door once finished. On the control program, insure that the “Interlock” indicator is yellow once the door is closed to ensure that the machine will turn on.

8. Active the x-rays. Press the large “On” button on the control program. The red light on top of the machine will turn on, to indicate that the x-rays are active.

9. Open the viewer. On the large “Image” window, press the “Live” button, which resembles a video camera. The button is the 5th from the left in the toolbar. A smaller window will open which displays the active image.
10. Orient the sample to the region that is to be imaged. The $x$- and $y$-axis positions of the stage and $z$-axis control of the x-ray source can be controlled with the gaming controller next to the machine. Once the sample is positioned in the general area by looking through the window, the live x-ray image on the screen can be used to orient the device.

11. Capture the image. Once the desired position has been found, on the large “Image” window, press the “Capture” button, which resembles a rectangle with a circle in the middle. The button is the 6th from the left in the toolbar. The image will momentarily disappear and reform with higher resolution.

12. Process the image. A number of image processing filters are available in the software. There are located in the large “Image” window by selecting “Image” $\rightarrow$ “Filter”, then selecting the chosen filter. The “Enhance Edge” filter is particularly useful.

13. Save the image. Save the image to the computer within the “Image” window by selecting “File” $\rightarrow$ “Save”.

14. Repeat the process for additional images.

15. Deactivate the x-rays. Press the large “Off” button on the control program. The red light on top of the machine will turn off, to indicate that the x-rays are no longer active.

16. Unload the sample.

17. Close the programs and save the images to an external drive. Close both the control and viewer programs by clicking the “X” in the upper right corner. Attach an external drive and transfer the images.

18. Turn off the equipment. Repeat the procedure of step 1 in reverse. First, turn off the computer using the start menu. Wait for the computer to power off. Second, flip the switch next to the key to the “Off” symbol. Third, turn the key.

19. Complete an entry in the log book. The log book is a red binder on top of the equipment. Please complete all fields so that a record of use is maintained.

Note: This process can be conducted both after die attachment and/or after vacuum sealing. It is useful to observe the solder reflow and void presence after both attachment procedures.
C.3 Breadboard Front-End Electronics

This is a step-by-step guide to creating breadboard front-end electronics for the electrostatic actuation and detection of MEMS, specifically using electromechanical amplitude modulation. The following items are required: Packaged MEMS resonator for testing, socket for interfacing package to breadboard, breadboard, operational amplifiers (AD620 x 3), instrumentation amplifiers (OP177 x 2), resistors (500kΩ x 3, 10kΩ x 2, 1kΩ x 2), capacitors (1µF x 2, 10pF x 2), jumper wires, power supplies (±15 Volts and second for controlling DC voltages (5 Volts)), lock-in amplifier, spectrum analyzer, BNC cables (6), BNC splitter (1), BNC to wire connections (4), banana cables (4), banana to BNC connection (1), and optional oscilloscope for debugging. For additional details or specifications, please see Appendix A.

Actuation will be accomplished by mixing AC and DC voltages on the breadboard and applying this new signal to the MEMS device. The AC voltage will be delivered by the spectrum analyzer and the DC voltage will be delivered by an independent power source. This actuation signal can then be applied to any number of electrodes, including differentially (separated by 180 degrees). The applied voltage creates a force on the resonant structure, potentially driving it into resonance.

For detection, electromechanical amplitude modulation will be used, which involves applying an AC signal to the proof mass of the structure with zero DC offset. This is known as the carrier signal, because it carries information about the motion of the structure when detected, and is supplied by the lock-in amplifier. This signal passes over a variable gap (the value of which is based on the displacement of the resonator) and is then detected using one or multiple detection electrodes (potentially differential). The electrodes output a current which is a function of both the carrier signal and motion of the resonator. This current will then be converted to a voltage and added together on the breadboard from any number of electrodes. The voltage output is then fed back into the lock-in amplifier, which extracts only the motion of the resonator, and into the spectrum analyzer for detection as a function of changing the AC actuation frequency.

C.3.1 Build Circuit on Breadboard

1. Collect the following electrical components:
   
   (a) Breadboard
   (b) Socket for interfacing package to breadboard
   (c) Operational amplifiers (AD620 x 3)
   (d) Instrumentation amplifiers (OP177 x 2)
   (e) Resistors (500kΩ x 3, 10kΩ x 2, 1kΩ x 2)
   (f) Capacitors (1µF x 2, 10pF x 2)
   (g) Copper Wire for jumper wires (with snips and wire stripper)

2. Connect rails to eventually become 15 Volts and ground. Figures C.11 shows the basic layout of the circuit on a breadboard. Because each amplifier requires 15
Volts and ground, the top two rails are connected to ground (green) and -15 Volts (grey), while the bottom two rails are connected to +15 Volts (red) and ground (green). Note that the left and right halves of the horizontal rails are typically not connected on standard breadboards.

3. Insert amplifiers, leaving room for passive components and the socket. The breadboard of Figure C.11 is broken up into four segments: Actuation, Socket, Detection X, and Detection Y. Actuation mixes the DC and AC driving voltages. The Socket contains the MEMS. Because gyroscopes are usually tested, two identical detection circuits are shown in Figure C.11, one for each axis: Detection X and Detection Y. The components listed above are only for a single detection channel. Figures C.12 and C.13 show detailed schematics of the actuation and detection circuitry: actuation requires two instrumentation amplifiers, while detection requires two operational amplifiers and one instrumentation amplifier.

4. Connect each amplifier to the 15 Volts and ground rails, as shown in Figure C.11. If it is difficult to tell which pin is connected, please refer to the datasheet of the amplifier in question. In Figure C.11, the wires are color coded: +15 Volts (red), -15 Volts (grey), Ground (green).

5. Connect passive components and wiring. The yellow wires of Figure C.11 show the internal connections for the drive and detection circuitry. These wires, along with the resistors and capacitors, can then be connected as outlined in Figures C.12 and C.13.

Note: The components listed here assumes a parasitic capacitance of the MEMS device to be 10pF, with resonance frequencies anywhere from 0.5 to 25 kHz. Beyond these ranges, new electronics will need to be developed.
Figure C.12: Actuation circuit (IA = AD620, \( R_G = 500\,\Omega \), \( R_1 = 1\,\Omega \), \( C_1 = 1\,\mu\text{F} \)).

Figure C.13: Detection circuit (OA = OP177, IA = AD620, \( R_G = 500\,\Omega \), \( R_F = 10\,\Omega \), and \( C_F = 10\,\text{pF} \)).
C.3.2 Build Test Setup

1. Collect the following equipment:
   
   (a) Power supplies
       i. BK Precision 1672 Triple Output DC Power Supply (to supply 15 Volts)
       ii. Agilent E3612A DC Power Supply (to supply VDC)
   (b) Ametek Signal Recovery Model 7270 DSP Lock-in Amplifier
   (c) Agilent 35670A Dynamic Signal Spectrum Analyzer
   (d) Oscilloscope and multi-meter (optional for debugging)
   (e) Cables and connectors
       i. BNC cables (6)
       ii. BNC splitter (1)
       iii. BNC to wire connections (4) (may require soldering to fabricate)
       iv. Banana cables (4)
       v. Banana to BNC connection (1)

2. If necessary, create the BNC to wire connections. This may involve soldering and will result in a BNC connector input with two wires as output: signal and grounding. Signal is connected to the center pin, while grounding is used to surround the signal to reduce noise. Four of these connectors are required.

3. Use three of the BNC to wire connectors to connect to the breadboard for VAC, VDC and VO. The fourth one is for the carrier signal and will be used later. The locations on the breadboard for each of these connections are shown in Figure C.14 for the example breadboard, and schematically in Figures C.12 and C.13. The 'signal' wire to the BNC must be connected in these locations, with the 'grounding' wire connected to the ground rail on the circuit.

4. There are two variable voltages sources on the BK Precision power supply with three ports each: “-”, “G”, and “+”. Connect a banana cable from the “+” port on the left source to the “-” port on the right source.

5. Connect a banana cable from the “-” port on the left source to the -15 Volt rail on the breadboard.

6. Connect a banana cable from the “+” port on the right source to the +15 Volt rail on the breadboard.

7. Connect a banana cable from the “-” port on the right source (or the “+” port on the left source) to the ground rail on the breadboard.

8. Connect the banana to BNC connector to the “+” and “-” ports of the Agilent power supply. Make sure the positive port is connected to the center pin of the BNC connection by inserting the connector so that the asymmetric lump is facing upwards. Double check with multi-meter.
9. Connect a BNC cable from the banana to BNC connector to the BNC to wire connection attached to VDC on the breadboard.

10. Attach the BNC splitter to the “Source” port on the back of the spectrum analyzer.

11. Connect a BNC cable from one of the two “Source” connections on the back of the spectrum analyzer to “Ch 1” on the front of the spectrum analyzer.

12. Connect a BNC cable from the second “Source” connection on the back of the spectrum analyzer to VAC on the breadboard.

13. Connect a BNC cable from “Ch 2” on the front of the spectrum analyzer to DAC 1 on the back of the lock-in amplifier.

14. Connect a BNC cable from “A” on the front of the lock-in amplifier and connect it to the BNC to wire connection attached to VO on the breadboard.

15. Connect a BNC cable from “OSC Out” on the front of the lock-in amplifier to the last BNC to wire connection that was never attached to the breadboard.

16. Double check your setup so that it looks similar to Figure C.15.

17. On both power supplies (BK Precision and Agilent), turn each of the voltage and current knobs as far counter-clockwise as they will go.

18. Turn on the BK Precision power supply. It should read 0 Volts and an LED indicator should say that the output is current limited.

19. On one of the channels, slowly turn both the voltage and current knobs clockwise until a positive voltage is reached and the LED indicator says the output is voltage limited. Increase voltage until either a value of 12 Volts is reached or the LED indicator switches to saying that the output is current limited. When it says it is current limited, increase the current knob slightly until it returns to being voltage limited and repeat until 12 Volts is reached. Essentially, an output of 12 Volts is desired; however, current should be limited to only what is necessary to power the amplifiers. When current is not limited, there is a risk of damaging equipment or the MEMS.

20. Repeat steps 18 and 19 for the second BK Precision voltage channel, again setting the voltage to 12 Volts.

21. Repeat steps 18 and 19 a third time for the Agilent power supply, this time setting the voltage to 5 Volts, or whatever DC voltage that will be use to drive the device.

22. Turn off both power supplies.

Note: Different makes and models of this equipment may be adequate, but this guide is written for the specific types mentioned above.
Figure C.14: Interface locations for BNCs ($V_{AC}$, $V_{DC}$, $V_O$) and ports for jumper wires to the socket ($V_{A1}$, $V_{A2}$, $V_{D1}$, $V_{D2}$).

Figure C.15: Basic set-up of equipment and connections.
C.3.3 Interface Package on Breadboard

1. Determine the electrode pin locations. For the specific MEMS being tested, wire-bonding has already been completed and the interfacing pins determined. Using this scheme, the pins of the packages which correspond to the following electrical connections to the MEMS should be known. If a gyroscope is being tested, there should be two sets of pins B through E: one of the X-axis and one for the Y-axis.

   (a) Proof mass
   (b) Drive +
   (c) Drive -
   (d) Sense +
   (e) Sense -

2. Using jumper wires, connect the actuation and detection electronics to the pins on the breadboards that correspond to the correct pins of the package. Drive + corresponds to $V_{A1}$, Drive - corresponds to $V_{A2}$, Sense + corresponds to $V_{D1}$, and Sense - corresponds to $V_{D2}$. For single-sided detection, $V_{A2}$ and $V_{D2}$ are not needed. Figures C.12 through C.14 can be used as a guide.

3. Connect the last BNC to wire connector to the breadboard, where the BNC connector is attached to the “OSC Out” port on the lock-in amplifier. The “signal” wire should be connected to the proof mass of the MEMS structure and the “ground” on the ground rail.

Note: Some MEMS wirebonding schemes also have connects that must be connected to the ground rail (such as the substrate of the device). If these connections need to be made, they should be denoted in the wirebonding scheme.

C.3.4 Modify Lock-in Amplifier Settings

1. Turn on the lock-in amplifier by switching the power switch on the back. Wait for it to start up.

2. Set the “DAC 1” output to “X%”. Press Menu → Main menu 2 → DAC Menu. The value for “DAC1 SETUP” should be set to “X% (2.5V fs)”. Press the Main Display button to exit.

3. Set the Oscillator frequency to 52 kHz and amplitude of 0.5-2 Volts. Press Menu → Oscillator. “OSC FREQUENCY” should be set to “52,000.00 Hz” and “OSC AMPLITUDE” should be set to “2,000 000 V”. Press the Main Display button to exit.

4. The main display screen should display “R” and “Θ” on the right in a large font. If it does not, press the buttons next to each until the top value is set to “R” and the bottom value is set to “Θ”.

415
5. Set “AC Gain” to “0dB, DR 20” on the main display.
6. A green light at the top of the display should read “LOCKED”.

Note:

C.3.5 Modify Spectrum Analyzer Settings

1. Turn on the spectrum analyzer by pressing in the “POWER” button in the lower left corner of the front panel. A green light should turn on. Wait for it to start up.
2. Press “Inst Mode” → “SWEPT SINE”.
3. Press “Meas Data” → “FREQ RESP 2/1”.
4. Press “Trace Coord” → “dB MAGNITUDE”.
5. Press “Scale” → “Y PER DIV (DECADES)”.
6. Press “Scale” → “Autoscale ON”.
7. Press “Active Trace” → “A”.
10. Set the settling and integration time. Settling time determined how much time passes between actuation and measurement. Increasing settling time can help with sweeping high-Q devices. Integration time determined the number of averages that is collected for each measurement. Press “Avg” → “SETTLE TIME” → “2” → “0” → “CYCLES” for 20 frequency cycles between actuation and measurement. Press “Avg” → “INTEGRATE TIME” → “2” → “0” → “CYCLES” for 20 averages for each data point. For a cleaner sweep, increase both of these values to 200.
11. Set the number of data points collected. Press “Freq” → “RESOLUTN SETUP” → “RESOLUTN” → “2” → “0” → “1” → “POINT / SWEEP” for 201 data points. Press “Rtn” to return to the main screen.
12. Set the frequency range of measurement. Press “Freq” and set the range by either selecting a center frequency and frequency span around it (by setting “SPAN” and “CENTER”), or by setting a start and stop frequency (by setting “START” and “STOP”). For low-frequency QMGs, this should be anywhere from 2kHz to 3.5kHz.

Note:
C.3.6 Sweep Device and Save Responses

1. Check that both DC power supplies are off.

2. Insert the MEMS into the socket.

3. Turn on both DC power supplies.

4. On the lock-in amplifier, change “SENSITIVITY” until the value of “R” is close to 100%. The “SENSITIVITY” value should be close to “200 mV”. If not, the parasitic capacitance of the device may be far from 10pC.

5. On the lock-in amplifier, press Menu → Auto functions → AUTO PHASE. The main display should not be seen on the screen and “Θ” should be close to zero.

6. On the spectrum analyzer, press “Start” to begin the frequency sweep. At any point, “Pause Cont” can be pressed to pause or continue the sweep, or “Start” to begin the sweep again.

7. Sweep the device multiple times until the frequency response is found. Typically, simply changing the frequency range will suffice; however, the driving force can also be manipulated by changing both the AC and DC driving voltages. Please note that for vacuum-sealed devices, substantially less driving voltage is required. For devices in air, AC voltage is set to 5 V (the maximum that the spectrum analyzer can supply) and DC voltage is also set to 5 V. For additional driving force, the DC voltage can be increased. For devices in vacuum, both AC and DC voltages should be decreased as a function of the vacuum level. Please note that DC should always be higher than the AC voltage. Driving device under vacuum with too great of a force can damage them irreversibly. Also, make sure you are properly limiting current on the DC power supplies before ever testing vacuum-sealed devices.

8. To help identify if a magnitude peak is truly a mechanical resonance, the phase of the response can also be examined. Press “Trace Coord” → “PHASE” to examine the phase response, which should shift by roughly 180 degrees for a mechanical response. The magnitude response can be redisplayed by pressing “Trace Coord” → “dB MAGNITUDE”.

9. To save the response once one has been found, first insert a 3.5 inch floppy on which to save the data.

10. Press “Save / Recall” → “SAVE DATA”. Check to make sure that the “FORMAT” is set to “SDF”. Continue by pressing “SAVE TRACE” → “INTO FILE”. Change the file name to whatever is desired. Note there is a limit to the number of characters. Press “ENTER”. A green light on the floppy drive will turn on. Wait until this light turns off.

11. If desired, take additional sweeps, saving them as above under different names on the floppy. Note that you may not be given warning if you are about to save over existing data, so be careful with naming.
12. Remove the 3.5 inch floppy and insert it into a computer with a 32bit operating system and Matlab.

13. Upload the files saved to the floppy on the hard drive of the computer into a new folder.

14. Download the following programs from \NITRIDE\Public to the same folder as the data.
   (a) run.bat
   (b) Sdftoml.exe

15. Edit the file “run.bat”. For each saved plot, type the following line, where “name” is the name of the file.
   (a) SDFTOML.exe /x name.dat name.mat

16. Double click “run.bat” to run the program. This will create a new data file for each sweep which can be read by Matlab.

17. Open Matlab and go to “File” → “Import Data…”. Select one of the new data files and press “Open”. This will generate two arrays: o2i1x and o2i1. Note that the name “o2i1” can change.

18. Magnitude and phase can then be plotted using the following lines.
   (a) Magnitude: plot(o2i1x,20*log10(abs(o2i1)))
   (b) Phase: plot(o2i1x,phase(o2i1)*180/pi)

Note: If R cannot be set to approach 100% or if there is a blinking red light at the top that says “INPUT OVERLOAD”, there is likely an electrical short somewhere inside the package or in the circuit. For a device that has not been vacuum sealed yet, this is many times fixed by holding the package upside-down and tapping on the back hard with tweezers. In addition, the package can also be tapped upside-down on a hard, clean surface to try and remove the material causing the short.
C.4 Open-loop Gyroscope Operation using a Zurich HF2LI Lock-in Amplifier

This is a step-by-step guide to using the Zurich HF2LI Lock-in Amplifier for open-loop control of gyroscopes. You will need a single HF2LI, a nearby PC with USB port to control it, BNC cables, and a gyroscope mounted in front-end electronics (with all necessary power supplies).

C.4.1 Build Test Setup

1. Download and install Zurich HF2LI Lock-in Amplifier software.
   (a) Software can be found on their website: http://www.zhinst.com/products/hf2li
   (b) Or on Nitride: \\NITRIDE\Software

2. Download the default configuration file from \\NITRIDE\Public
   (a) !gyro.zicfg

3. Configure hardware:
   (a) Zurich HF2LI Lock-in Amplifier.
      i. Connect power cable on back of HF2LI.
      ii. Connect USB cable from “PC USB 2.0” on back of HF2LI to controlling computer.
      iii. Connect BNC from “X/R1 Aux 1” on front to “Aux In 2” on back of HF2LI.
   (b) Front-End Electronics of Gyroscope.
      i. Connect BNC from “Signal Input 1: +In” on front of HF2LI to output of sense mode.
      ii. Connect BNC from “Signal Input 2: +In” on front of HF2LI to output of drive mode.
      iii. Connect BNC from “Signal Output 1: Out” on front of HF2LI to proof mass.
      iv. Connect BNC from “Signal Output 2: Out” on front of HF2LI to drive mode excitation.
      v. Connect necessary power of front-end electronics.
      vi. Connect the DC excitation biases for both drive and sense mode to external power source: 2 Volts.

Note:
C.4.2 Load Default Program

1. Open the program: “ziControl”.

2. Turn on power switch of HF2LI on back of HF2LI, next to power cable.

3. Open the “Save” tab. Click on the “Load Settings” button and load the configuration file “gyro.zicfg”.

4. Open the “Modulation” tab. Make sure that both Mod1 and Mod2 are using Oscillator 3 as the “Carrier” and Oscillator 2 (PLL 2) as the “Modulation”. If not, try loading the configuration file again. The correct settings are shown in Figure C.16.

5. Open the “Auxiliary” tab. For Aux 1, click on the arrows pointing to one another between “Offset” and “Value” to zero the value.

6. Open the “Connectivity” tab. Under “Device Configuration”, the “Clock Source” should be set to “Internal Quartz”. Note: This program consists of two separate segments, each with their own independent set of tabs. The upper segment consists of the configuration and control settings, while the lower segment can display relevant information in several meaningful ways.

Note:

C.4.3 Establish Appropriate Carrier Frequency

1. Open the “Lock-in MF” tab on the top. This tab displays the eight demodulators of the system. From top to bottom, they represent:

   (a) The sense mode demodulation at the carrier frequency.
   (b) The sense mode demodulation at the first mechanical sideband.
   (c) The sense mode demodulation at the second mechanical sideband.
   (d) The drive mode demodulation at the carrier frequency.
(e) The drive mode demodulation at the first mechanical sideband.
(f) The drive mode demodulation at the second mechanical sideband.
(g) The demodulation of the oscillator controlled by PLL 1.
(h) The demodulation of the oscillator controlled by PLL 2, also used as drive
mode excitation.

2. In the “Lock-in MF” tab, the first demodulator will be fed to the proof mass
connected to “Output 1”, while the eighth demodulator will be fed to the drive
mode AC excitation connected to “Output 2”. Frequency of excitation can be
controlled under the “Demodulators” section and “Frequency” column. Amplitude
can be controlled under the “Signal Output Amplitudes” section and “Output 1”
and “Output 2”, along with turning on and off these excitations in the output
channels. The first demodulator should be currently exciting “Output 1”, and the
eighth demodulator should be currently exciting “Output 2”.

3. Open the “Oscilloscope” tab on the bottom. Under “Signal Input”, select “Signal
Input 2” to display the output of the drive mode of the sensor. This output should
mostly be the feedthrough of the carrier frequency.

4. In the “Lock-in MF” tab, adjust the amplitude of “Output 1” of the first demod-
ulator until the oscilloscope appears sinusoidal. The value should be as high as
possible, while still being sinusoidal (without saturating amplifiers). An example
of saturated amplifiers is shown in Figure C.17, while a correct response is shown
in Figure C.18.

Note: The maximum magnitude and frequency of the carrier is largely determined by
the amplifiers in the front-end electronics.

C.4.4 Sweep Sense Mode to Confirm Resonance

1. Currently, the drive mode AC excitation of the sensor is physically connected to
“Signal Output 2” on the front of the HF2LI. To sweep the sense mode in order
to confirm resonance and characterize the frequency separation, disconnect “Signal
Output 2” from the drive mode AC excitation and reconnect it to the sense mode
AC excitation.

2. Open the “PLL” tab on the top. Under the large labels “PLL1” and “PLL2” there
are small buttons labels “En”. These buttons turn on and off their respective PLLs.
They should both be currently off; the buttons should be grey.

3. Open the “Sweeper” tab on the bottom. Under the “Sweep Range” section, enter
“Start” and “Stop” frequencies to sweep through. For QMGs, this can range from
2 kHz to 3.2 kHz. The sweeper tab is shown in Figure C.19.

4. In the “Sweeper” tab, under “Signal Input”, select “Input Demod” as “2”.

421
Figure C.17: Initial settings under the “Connectivity” tab, along with a saturated carrier frequency in the “Oscilloscope”.

Figure C.18: Initial settings under the “Connectivity” tab, along with an unsaturated carrier frequency.
5. In the “Sweeper” tab, under “Sweep Control”, select “Oscillator” as “2”.

6. Open the “Lock-in MF” tab on top. The eighth demodulator will be exciting the sense mode. Make sure the amplitude is a reasonable level depending on the Q-factor of the device. For low-Q devices, 0.1-1 Volt.

7. In the “Sweeper” tab, under “Sweep Control”, click “Single” to begin the sweep. In the windows to the left, the magnitude and phase of the response will be shown. When the sweep in complete, you should see some kind of a response in magnitude and phase at a specific frequency, though it might not be clear. If you do not see a response, either narrow the frequency range of interest or increase the amplitude of actuation.

8. At the lower left of the magnitude chart, select the zoom tool (only an icon with no text). This can be used to zoom into the frequency range of interest. Under “Sweep Range”, click “Copy from Plot Range” to update the “Start” and “Stop” frequencies with the zoomed into view.

9. Under “Sweep Control”, click “Single” again to sweep in the new range to get a more clear view of the sweep, as shown in Figure C.19. This can be repeated until an adequate resonance sweep is achieved.

10. In the lower right corner of the “Sweeper” tab, there are four additional tabs: Display, Cursor, Calibration, and Q-factor. Select “Cursor” to insert cursers on the plot. Drag one to the center of the peak magnitude to measure the resonance frequency. Record this number, along with Q-factor (in Q-factor tab).

Note:

C.4.5 Sweep Drive Mode to Find Resonance Frequency for PLL

1. Currently, the sense mode AC excitation of the sensor is physically connected to “Signal Output 2” on the front of the HF2LI. The drive mode must be replaced to continue. Disconnect “Signal Output 2” from the sense mode AC excitation and reconnect it to the drive mode AC excitation.

2. The steps from the previous section can be repeated to find the drive mode resonance frequency, with the following exception:

   (a) In the “Sweeper” tab, under “Signal Input”, select “Input Demod” as “5”. This is so that you will be detecting the drive mode output signal instead of the sense mode.

Note:
C.4.6 Establish Drive Mode Phase-Locked Loop (PLL)

1. In the “Sweeper” tab, click “Single” again to begin a sweep once more.

2. During the sweep, switch over to the “Spectroscope” tab and under “Control”, make sure “Demod 5” is on. You may also have to adjust the “Time Scale” to view a larger time range.

3. As the sweep continues, the output should rise and fall, as shown in Figure C.20. Once you see the peak amplitude, under “Control”, click “Acq Stop” to pause the measurement.

4. Under “Cursors”, turn “C1” on and drag it to the peak of the response. Under “Cursors” the “Value” will display the magnitude at that location. Remember or record this number and click “Acq Stop” again to continue reading the output.

5. Open the “PLL” tab on top. “PLL2” will be used to excite the drive mode. Under the “PLL Setpoint” section, type the drive mode resonance frequency into the “Center F” box of “PLL2”. “Range” should be around 12 Hz and represents the frequency bandwidth around the center frequency for which the PLL will look for the resonance frequency of the device. “Set point” should be +180 degrees.

6. Under the “PLL2” label, click the “En” button to enable the PLL.
7. Under “PLL Setpoint”, reduce the “Set point” of “PLL2” by 10 degrees at a time and observe the output of the Spectroscope. The value should start to rise at some point. Continuing adjusting the set point until the output of the Spectroscope is similar to the value observed earlier in this section. This process is shown in Figure C.21.

8. The PLL should now be locked. To confirm, in the “PLL Setpoint” section, “F shift” should now be near zero. You can also double check by opening the “Lock-in MF” tab. There should now be a green light to the right of the eighth demodulation that says “Locked”. The drive mode will now continue to resonate with the set AC excitation voltage, even if frequency of the device shifts, but not if it shifts beyond the “Range”.

Note:

C.4.7 Establish Drive Mode Amplitude Gain Control (AGC)

1. Open the “PID 1” tab. Under “Input”, the “Set” value is the set-point for the AGC, or the voltage value you desire the output will stay at. This output should be similar to the AC excitation voltage you have been using up to this point.

2. In the “PID1” tab, in the “Output” section, the range of AC excitation voltage that the AGC is allowed to use is dictated by the “Center” value, plus/minus the
Figure C.21: Spectroscope data as the phase “Set Point” is adjusted for PLL2 until maximum value is reached.

“Range” value. This is to establish limits to avoid breaking the device should an error occur.

3. In the “PID 1” tab, in the “Control” section, the PID gains of the controller can be manually adjusted. These gains can also be automatically tuned using the “DUT System Model” by inserting the frequency and Q-factor of the resonator and clicking “Tune”.

4. In the “PID 1” tab, under “Control”, click the button next to “Enable” to enable the AGC. The output shown in the spectroscope will continually attempt to achieve the “Set” value under “Input”, using the range of excitation voltages it is allowed.

5. Test the AGC by observing the output of the Spectroscope. Change the “Input” value of the AGC and watch the Spectroscope output slowly approach this value.

Note:

C.4.8 Record Data on Bias

1. Both PLL and AGC should currently be active.

2. Open the “Lock-in MF” tab on top. In the second demodulator, click the arrow to the right of “Phaseshift” to optimize the signal.
3. Open the “Spectroscope” tab on the bottom. In the “Control” section, turn off “Demod 5” and turn on “Demod 2” to observe the output of the sense mode of the device. You are now observing the output signal.

4. Open the “Save” tab on the top. Select “Demod 2” and “Demod 5” to save the output of both the drive and sense channels of the device.

5. In the “Save” tab, type in the location you want to save the data and click the “Save” button to begin recording data. After a few minutes (or more or less), click the “Save” button again to stop recorded data. This data can now be analyzed using Allan Variance to determine the white noise, bias instability, and rate random walk.

Note:

C.4.9 Record Data on Scale Factor

1. To measure scale factor, the gyroscope must be on a rate table. Please be cautious of cable interference whenever you start the rate table.

2. Give the rate table a sinusoidal rotation and observe the response of “Demod 2” in the Spectroscope. You should see the sinusoidal rotation, as shown in Figure C.22.

3. Open the “Lock-in MF” tab on top. In the second demodulator, click the arrow to the right of “Phaseshift” to optimize the signal.

4. Output data can be recorded in the same way as bias for a series of rotational rates to determine scale factor.

Note:
Figure C.22: Spectroscope data as the gyroscope is given a sinusoidal input rotation from a rate table.
Appendix D

Vendors

Below is a list of vendors whose services have been used during the course of this work, along with their contributions. They are listed alphabetically by vendor name.

Accretech
Website: http://accretechamerica.com/
Point of Contact: Keith Kruger, <krugerk@accretechamerica.com>
Phone: (510)344-5411 ext. 2412
Laser dicing of wafers. It is a dry, clean process that does not require front-side sensor protection, but can not cut through metal films.

American Precision Dicing, Inc.
Website: http://wafer-dicing.com/
Point of Contact: Remy Daniels, <remy.daniels@wafer-dicing.com>
Phone: (408)254-1600
Mechanical dicing of wafers. A wet process that requires front-side sensor protection. Can cut through nearly any MEMS material.

Capitol Scientific
Website: http://capitolscientific.com/
Point of Contact: Misty Hull, <misty.hull@capitolscientific.com>
Phone: (800)580-1167
Supplier of photoresist and developer for lithography.

Digi-Key Corporation
Website: http://www.digikey.com/
Point of Contact: N/A
Phone: (800)344-4539
Supplier of electronic components, useful for front-end interfacing of capacitive MEMS.
GGB Industries, Inc.
Website: http://www.ggb.com/
Point of Contact: Matt Birmingham, <matt@ggb.com>
Phone: (239)643-4400

Manufacturer of wedge cards for electrical probing of dies with fixed bond pad locations.

Grinding and Dicing Services, Inc.
Website: http://www.wafergrind.com/
Point of Contact: Jeremy Favia, <JFavia@wafergrind.com>
Phone: (707)508-9726

Wafer grinding for thinning and polishing silicon wafers.

Ideal Aerosmith, Inc.
Website: http://www.ideal-aerosmith.com/
Point of Contact: Jason Eder, <JEder@idealaero.com>
Phone: (218)773-5044

Supplier of precision rate tables for inertial sensor characterization.

Idonus Sarl
Website: http://www.idonus.com/
Point of Contact: Christian Spoerl, <christian.spoerl@idonus.com>
Phone: +41 32 724 44 40

Supplier of equipment for vapor-phase HF etching and infrared imaging of silicon wafers and dies.

Innovative Micro Technology
Website: http://www.imtmems.com/
Point of Contact: Suresh Sampath <Suresh@imtmems.com>
Phone: N/A

MEMS foundry for device fabrication.

InterMEMS
Website: http://www.intermems.com/
Point of Contact: Wendell McCulley <wmcculley@earthlink.net>
Phone: (408)241-0007

Wafer bonding of silicon and glass wafers.
Invensense
Website: http://www.invensense.com/
Point of Contact: Julius Tsai <jtsai@invensense.com>
Phone: N/A
Commercial MEMS manufacturer and MEMS foundry for device fabrication with on-chip electronics.

iX-Factory
Website: http://ix-factory.de/
Point of Contact: Hans Bouwes <H.Bouwes@ix-factory.de>
Phone: +49 231 47730 580
MEMS foundry for device fabrication.

Kimball Physics
Website: http://www.kimballphysics.com/
Point of Contact: N/A
Phone: (603)878-1616
Supplier of vacuum chambers.

Kurt J. Lesker Company
Website: http://www.lesker.com/
Point of Contact: N/A
Phone: (800)245-1656
Supplier of vacuum pumps, hoses, and miscellaneous vacuum supplies, along with evaporation materials.

L-Com
Website: http://www.l-com.com/
Point of Contact: N/A
Phone: (800)341-5266
Supplier of electronic cabling.

Marco Rubber
Website: http://www.marcorubber.com/
Point of Contact: Donna King, <donna@marcorubber.com>
Phone: (603)468-3600 ext. 3606
Supplier of custom o-rings.
Mark Optics, Inc.
Website: http://www.markoptics.com/
Point of Contact: Lily Sandoval <Lily@markoptics.com>
Phone: (714)545-6684
Supplier of glass wafers.

Materion Microelectronics & Services
Website: http://www.materion.com/
Point of Contact: Elizabeth Cattarin <elizabeth.cattarin@markoptics.com>
Phone: (716)837-1000
Supplier of 80/20 Au/Sn solder preforms for die attachment and lid sealing.

McMaster-Carr
Website: http://www.mcmaster.com/
Point of Contact: N/A
Phone: (562)692-5911
Supplier of miscellaneous hardware components.

MEMSCAP
Website: http://www.memscapinc.com/
Point of Contact: Buzz Hardy <buzz.hardy@memscapinc.com>
Phone: (919)248-1486
MEMS foundry for device fabrication.

MetricTest
Website: http://www.metritictest.com/
Point of Contact: N/A
Phone: (800)432-3424
Supplier of electronic test equipment.

Mouser Electronics
Website: http://www.mouser.com/
Point of Contact: N/A
Phone: (800)346-6873
Supplier of electronic components, useful for front-end interfacing of capacitive MEMS.
Photo Sciences, Inc.
Website: http://www.photo-sciences.com
Point of Contact: Kiomi Hamada <kiomih@photo-sciences.com>
Phone: (310)634-1500 ext. 1524

Manufacturer of custom lithography masks.

Plastronics, Inc.
Website: http://www.plastronics.com/
Point of Contact: N/A
Phone: N/A

Supplier of sockets for interfacing packaged MEMS, specifically packages for the 44-pin LLC packages (Part Number: P2044S-B-Au).

Precision Instruments
Website: http://www.precisioninstrumentsvacuum.com/
Point of Contact: Margaret Hansen, <sales@precisioninstrumentsvacuum.com>
Phone: (714)630-9322

Local repair service for vacuum pumps. They are located in Anaheim, CA.

SAES Getters USA, Inc.
Website: http://www.saes-group.com/
Point of Contact: Heather A Florence <Heather_Florence@saes-group.com>
Phone: (719)527-4116

Getter material deposition for achieving $\mu$ - Torr vacuum levels in packaged cavities.

Semiconductor Packaging Materials, Inc.
Website: http://www.sempck.com/
Point of Contact: Dave Virissimo, <dvirissimo@sempck.com>
Phone: (619)464-5430

Supplier of wire bonding wire and custom Au/Sn solder preforms.

Semiprobe
Website: http://www.semiprobe.com/
Point of Contact: Denis Place <denis@semiprobe.com>
Phone: (802)860-7000

Manufacturer of systems capable of electrical probing of dies under vacuum.
Spectrum Micromechanical, Inc.
Website: http://www.spmmi.com/
Point of Contact: Jeff Olson, <jeff@spmmi.com>
Phone: (858)395-2264

Mechanical dicing of wafers. A wet process that requires front-side sensor protection. Can cut through nearly any MEMS material. They were typically used for dicing glass lids for vacuum packaging.

Spectrum Semiconductor Materials
Website: http://www.spectrum-semi.com/
Point of Contact: Mike Krulee <mwkrulee@spectrum-semi.com>
Phone: (408)435-5555

Supplier of standard and custom MEMS packages, specifically: Deep 44-LCC packages (Part Number: LCC04438), Shallow 44-LCC packages (Part Number: LCC04420), LCC combo lids (Part Number: CL-62007), 24-DIP packages (Part Number: KD-78516-D-03), and DIP combo lids (Part Number: CL-55018).

SST International
Website: http://www.sstinternational.com/
Point of Contact: Paul W. Barnes, <pwbarnes@sstinternational.com>
Phone: (562)803-3361

Supplier of vacuum packaging equipment, services, and custom graphite machining.

Ted Pella, Inc.
Website: http://www.tedpella.com/
Point of Contact: N/A
Phone: (530)243-2200

Supplier of high-friction device storage containers (Part Number: AD-23T-00-X4) and SEM pin mounts.

Teledyne Scientific & Imaging, LLC
Website: http://www.teledyne.com/
Point of Contact: N/A
Phone: (805)373-4545

MEMS foundry for device fabrication.
Tiffany Associates
Website: http://www.tiffanyassociates.com/
Point of Contact: Randy Tiffany, <randy@tiffanyassociates.com>
Phone: (949)830-3577
Supplier of Signatone microprobe test equipment.

Transene Company, Inc.
Website: http://www.transene.com/
Point of Contact: Christopher Christuk, <sales@transene.com>
Phone: (978)777-7860
Supplier of gold and chrome etchant.

Ultrasil Corporation
Website: http://www.ultrasil.com/
Point of Contact: Raymond Martin Duque <rduque@ultrasil.com>
Phone: N/A
Supplier of custom silicon and silicon-on-oxide wafers.

Universal Temperature Processes
Website: http://www.unitemp.de/
Point of Contact: Astrid Birkner, <astrid.birkner@unitemp.de>
Phone: +49(0)8441 787663
Supplier and repair of vacuum furnaces.

West-Bond, Inc.
Website: http://www.westbond.com/
Point of Contact: Brendon King
Phone: (714)978-1551
Supplier of wire bonders and new bond tools.

Zurich Instruments AG
Website: http://www.zhinst.com/
Point of Contact: Flavio Heer, <flavio.heer@zhinst.com>
Phone: +41-44-5150415
Supplier of specialized lock-in amplifiers.
Appendix E

Lithography Masks

Below is a list of the primary sensor lithography masks that were used over the course of this work, along with the disbursement of sensor designs. The purpose of this information is to serve as a reference, should these masks be used in the future. This is not a comprehensive list of all fabricated lithography masks, but simply the masks with multiple sensor designs, which could lead to confusion during fabrication.

“Accel1” - FM Accelerometer Mask 1
As shown in Figure E.1, this mask defines the etching pattern for three different designs of the linear frequency modulated accelerometer, as outlined below.

Blue - “Cell0_LF” - Low Frequency FM Accelerometer (∼ 500 Hz Anti-Phase)
Green - “Cell0_HF” - High Frequency FM Accelerometer (∼ 2.5 kHz Anti-Phase)
Yellow - “Cell0_VHF” - Very High Frequency FM Accelerometer (∼ 5 kHz Anti-Phase)

“DieAttach1” - Die Attachment Definition 1
As shown in Figure E.2, this mask consists of 13 different sizes of squares, for the purpose of patterning the backside metallization of the SOI wafers. These sizes range from 170 µm to 8100 µm.

“DieAttach2” - Die Attachment Definition 2
As shown in Figure E.3, this mask consists of a single die attachment pattern: a 2 mm diameter circle. This mask is the standard backside production mask, as determined by the previous mask, “DieAttach1”.

“iX” - Commercial Fabrication Run at iX-Factory 1
As shown in Figure E.4, this mask design was outsourced for fabrication at iX-Factory, a MEMS foundry in Germany. This mask does not exist at UCI, but the design was sent to the vendor for fabrication at their facility. This mask contained four designs, as outlined below.

Green - Phase 2 QMG
Blue - Phase 1 QMG
Figure E.1: “Accel1” - FM Accelerometer Mask 1.

Figure E.2: “DieAttach1” - Die Attachment Definition 1.
Figure E.3: “DieAttach2” - Die Attachment Definition 2.

Figure E.4: “iX” - Commercial Fabrication Run at iX-Factory 1.
As shown in Figure E.5, with additional labeling provided in Figure E.6, this mask consisted of 12 different gyroscope designs and one accelerometer. The primary purpose of this mask was to experiment with a number of different alternations from the previous “Mask12”, which had been designed at UCI’s MicroSystems Laboratory.

“Mask14” - Gyroscope Mask 14
As shown in Figure E.7, this mask focused upon three different gyroscope designs, as determined from the successes of “Mask13”. This mask mainly focused the tradeoffs between frequency and sense capacitance, as well as the electrode design.

Yellow - “Cell0_New” - A high frequency QMG with additional capacitance
Blue - “QMG_new_sensecenter” - Phase 1 QMG
Green - “QMG_new_drivecenter” - Phase 1 QMG, with the location of the drive and sense electrodes swapped
Figure E.6: “Mask13” - Gyroscope Mask 13 - Labels.

Figure E.7: “Mask14” - Gyroscope Mask 14.
As shown in Figure E.8, this mask focused upon the Phase 1 QMG for producing a high volume of sensors. The high frequency design was also included in a limited capacity for further testing, as well.

Blue - “Cell0_New” - A high frequency QMG with additional capacitance
Green - “Cell0_new_sensecenter” - Phase 1 QMG

As shown in Figure E.9, a number of new design elements were tested for the eventual Phase 2 design of the QMG.

Blue - Phase 2 QMG
Red - Phase 2 QMG with single-sided electrodes
Purple - Test chips for passive quadrature canceling electrodes
Yellow - Phase 2 QMG with passive quadrature canceling electrodes

As shown in Figure E.10, this mask was used for producing a high volume of Phase 2 QMG sensors, consisting only of this design.
Figure E.9: “Mask16” - Gyroscope Mask 16.

Figure E.10: “Mask17” - Gyroscope Mask 17.