A High-Performance Tunneling Accelerometer

Edward Boyden
Osamah El Rifai
Brian Hubert
Maurice Karpman
Dave Roberts

Term Project
6.777, Introduction to Microelectromechanical Systems
Spring 1999, Prof. Stephen D. Senturia

TABLE OF CONTENTS

I. INTRODUCTION 4

I.1. Problem Statement 4
I.2. Scanning Probe Microscopy 4
I.2.i. Tunneling 5
I.2.ii. Current, applied bias, and tip-electrode separation 5

II. DESIGN METHODOLOGY 8

III. PROOF MASS AND MECHANICAL DESIGN 9

III.1. General Discussion of Accelerometer Geometry and Specifications 9
III.1.i. Accelerometer Sensitivity 9
III.1.ii. Accelerometer Response Time 10
III.1.iii. Cross-axis sensitivity 11
III.2. Overview of Cantilevered Beam Design 11

IV. SYSTEM MODEL 13

IV.1. Beam Flexural Dynamics 13
IV.2. Squeeze Film Damping 15
IV.3. Electrostatic Actuators 15
IV.4. Models for the feedback electronics 15
IV.5. Beam Design: Dynamic Considerations 16
IV.6. Location of Electrostatic Pads 17
IV.7. Vacuum operation 17
IV.8. Linearization and tunneling set point: 17
IV.9. Controller Design 17
IV.10. Simulation Results 20

V. ELECTRONICS 28

V.1. Design Principles 28
V.2. Single-conductor Design 28
V.3. Instrumentation: elements 29
V.4. Instrumentation: model 31
V.5. Instrumentation: implementation 33

VI. PROCESS FLOW 34

VI. 1. Overview 34
VI.2. The Process Flow 35

VII. CONCLUSION 43

APPENDIX A 45

A.1. Behavior of a Cantilevered Beam 45
A.2. Cross-Axis Sensitivity 46
A.3 Tolerance Analysis 47
A.4 Detailed geometric considerations 49

APPENDIX B 51

B.1. OVERVIEW 51
B.2. THE PROCESS FLOW 52
REFERENCES 57
BIBLIOGRAPHY 58

I. Introduction

I.1. Problem Statement

The goal of this project was to design an accelerometer based on the operating principles of Scanning Probe Microscopy (SPM). The process was constrained to form the tip on the same wafer as the proof mass. The performance targets for the device were as follows:

Full range	5g
Response time	10ms
Linearity	<1%
Cross-axis sensitivity	<1%
Self-test output	2.5g
Shock survivability	100ms, 100g shock with 1ms risetime

I.2. Scanning Probe Microscopy

A number of techniques have been developed which allow surfaces to be imaged with sub-angstrom resolution. These techniques all involve a tip on the end of a cantilever which is brought into close enough proximity to a surface to interact with the surface. Measurement of the interaction allows determination of the height of the surface at each point. The tip is scanned over the surface to create an image of the surface. These techniques are collectively known as Scanning Probe Microscopy (SPM) modes (How93). The sensitivity of SPM modes to deflection of the cantilever makes them excellent candidates for the fabrication of high sensitivity accelerometers.

The first decision the team faced was selecting the SPM mode for this project. The modes can be grouped by the method used to determine the vertical position of the tip. Three detection methods are in use; piezoresistance, optical and tunneling current.

Position detection using piezoresistance is accomplished by fabricating a piezoresistor at the root of the cantilever and measuring the change in resistance as the cantilever is deflected. This method is relatively poor at measuring small deflections and therefore yields relatively low resolution. Moreover, piezoresistive elements generally have complex variations with temperature. For these reasons, piezoresistive techniques are not appropriate for construction of a high sensitivity accelerometer.

The second, and most common, method of determining the vertical position of the tip is optical. A laser beam is bounced off the back of the cantilever onto a position sensitive photodetector. Deflection of the cantilever causes the position of the beam on the detector to shift. The ratio of the distance from the cantilever to the detector to the length of the cantilever amplifies the motion of the beam. As a result, this method can achieve sub-angstrom resolution.

The optical method has several advantages; it is extremely sensitive, it does not require a vacuum, it can be used with conductive and insulating samples, and it can be used with a variety of SPM modes. However, the need for a laser diode and a photodetector, and the size penalty imposed by the optical method make it undesirable for an SPM accelerometer.

The third method used to determine the vertical position of the tip is to measure the tunneling current between the tip and the substrate. This was the first SPM mode developed. Because the tunneling current is exponentially dependent on the separation between the tip the substrate, the distance between the tip the substrate can be measured to within 0.01 Angstrom. Despite the sensitivity of tunneling, it has been largely displaced by optical methods because it requires a vacuum for imaging oxide forming samples and can only be used with conductive tips and samples. These drawbacks are not serious impediments to the construction of an accelerometer. As a consequence, all of the SPM accelerometer work the team encountered in the literature relies on tunneling current as the detection method.

Based on the above logic and the existence proof of tunneling accelerometers in the literature, the team selected tunneling as the SPM mode for this project.

I.2.i. Tunneling

The proposed microelectromechanical (MEMS) tunneling accelerometer uses a general principle of operation that is commonly used for scanning tunneling microscopy (STM). In STM a bias voltage is applied between a sharp metal tip and a conducting sample. When the tip and sample are brought to within a few Angstroms (Å) of each other, a tunneling current can flow due to quantum mechanical tunneling effects. Because the tunneling current is exponentially dependent on the separation between the tip the sample, the distance between the tip the sample can be measured to within 10^{-3} (Å). Due to this high sensitivity to position, we propose that a MEMS cantilever structure with a variable gap between an integrated tunneling tip and a conducting electrode can be used in the construction of a tunneling accelerometer. Due to the excellent stability of its surfaces in vacuum or ambient gases, gold is the tunneling material of choice for both the tip and the electrode. This prevents drift in the observed tunneling current over time. However, platinum-iridium alloys with low reactivity can also be used to form the tip and the conducting electrode of the proposed tunneling accelerometer if very high fabrication temperatures limit the usefulness of gold.

I.2.ii. Current, applied bias, and tip-electrode separation

For the following discussion, we consider only the case of a one-dimensional metal-vacuum-metal tunneling junction. This is the configuration for the proposed tunneling accelerometer when operated within a vacuum instead of ambient gas. If the lateral dimension of the tip end is much larger than the characteristic wavelength related to the decay constant, $2\pi / \kappa \approx 6$ Å, which is expected for the proposed fabrication techniques used to create the tip, the one-dimensional treatment of tunneling parameters are approximately accurate (Che93, p. 7)

The following symbols are used in subsequent equations and will now be defined:

- E energy of the electron; typically measured in electron volts (eV), where 1 eV = 1.602177×10^{-19} J.
- V the applied voltage bias applied between the tip and the conducting electrode; in classical mechanics the electron may not penetrate into any potential barrier with $E < U(z)$
- m electron mass = 9.1×10^{-28} grams

$\hbar$ Planck constant over 2π; 6.582122×10^{-16} eV·s = 1.054572×10^{-34} J·s

ϕ effective local potential barrier height; also used interchangeably with the work function of a metal surface; usually refers to the conductive electrode and not the tip; defined as the minimum energy required to remove an electron from the bulk to the vacuum energy level; when measured in eV, the work functions of metals range from 4.1, 4.6, 4.8, and 4.8 for Al, Cu, W, and Si on the low end to 5.4, 5.6, and 5.7 for Au, Ir, and Pt on the high end [HCP].

Θ electron decay rate (in units of Å^{-1}); describes a state of the electron decaying in the positive z direction through the insulating barrier

z positive z is in a direction parallel to the "thickness" of the insulating barrier; $z = 0$ is located at the boundary of the metal sample and vacuum; $z = s$ is located at the boundary of vacuum and the metal of the tip

s the gap separating the tip from the conducting electrode; usually measured in the Angstroms; the gap is usually occupied by vacuum or ambient gas

Let us first consider tunneling between metal tip and the conducting electrode within the low-bias regime. Typically this means that the voltage between the tip and the conductive electrode is in the millivolt range. Depending on surface preparation, tip geometry, tunneling medium, and other factors, tunneling currents can range from nanoamps to microamps. In the low-bias regime, the vacuum tunnel junction exhibits Ohmic behavior so that tunneling current I is linearly proportional to the applied bias V. At the same time, tunneling current varies exponentially with size of the gap s between the tip and the plug. These relationships are illustrated in the following equations:

$$I \propto V \cdot e^{(-2s\Theta)} \tag{a}$$

where the electron decay rate Θ in units of Å^{-1} is obtained by

$$\Theta = \frac{\sqrt{2m\phi}}{\hbar} \approx 0.51\sqrt{\phi} \tag{b}$$

For example, if we assume a typical surface work function of 4 eV, the value for the decay rate Θ is about 1 Å^{-1}, and the current decays about $e^2 \approx 7.4$ times per Å of gap. The dependence of the logarithm of the tunneling current with respect to distance is a measure of the work function, or the tunneling barrier height. In fact, at *constant applied bias* V,

$$\phi[eV] = \frac{\hbar^2}{8m}\left(\frac{d\ln I}{ds}\right)^2 \approx 0.95 \cdot \left(\frac{d\ln I}{ds\left[\overset{o}{\text{A}}\right]}\right)^2 \tag{c}$$

As a consequence, we find

$$\left(\frac{d\ln I}{ds}\right) \approx const \tag{d}$$

In similar fashion, at *constant tunneling current*,

$$\phi[eV] = \frac{\hbar^2}{8m}\left(\frac{d\ln V}{ds}\right)^2 \approx 0.95 \cdot \left(\frac{d\ln V}{ds\left[\overset{o}{\text{A}}\right]}\right)^2 \quad \text{(e)}$$

Therefore, in the low-bias (millivolt) limit, where the tunnel junction exhibits Ohmic behavior, we simply find

$$\left(\frac{dI}{dV}\right) = \frac{I}{V} \quad \text{(f)}$$

Therefore, equation (a) reduces to

$$I \propto V \cdot e^{((const)(-s))} \quad \text{(g)}$$

because $\phi \approx const$ and $\Theta \approx const$. In summary, the applied bias voltage varies with the tip-electrode separation if the tunneling current is kept constant. For a constant tunneling current I, the quantity dI/dV therefore diverges like $1/V$ as V approaches zero. The problem of $1/V$ divergence can be solved by operating at a constant tunneling conductance $\sigma = I/V$ instead of constant tunneling current [Wie, p. 148]. For standard tunneling operations, tip-electrode distance ranges between 1 Å and 4 Å, which provides a corresponding conductance range of 10^{-6} to 10^{-8} Ω^{-1}. When taken to the logical conclusion, operation at constant conductance yields constant tip-sample separation distance s, as shown in the following derivation:

$$\sigma = \frac{I}{V} = const \propto e^{((const)(-s))} \quad \text{(h)}$$

$$\ln const \propto (-s) \cdot const \quad \text{(i)}$$

$$s = const \quad \text{(j)}$$

It should be noted that equations (h), (i), and (j) only hold in the low-bias regime when voltage and current and linearly related. For higher bias (on the order of volts) between the tip and the conducting electrode, the bias-dependence of the tunneling current generally does not exhibit Ohmic behavior. In the higher bias regime, the electronic states of both the tip and the electrode contribute to the tunneling current, and these states are dependent on the applied bias voltage. Tunneling current is proportional to a transmission coefficient, which itself is dependent on the bias. In general, the bias-dependence of the transmission coefficient at high-bias leads to an order-of-magnitude increase in the tunneling current for each volt increase in applied bias voltage (Wie94, p. 18).

In short, a tunneling accelerometer can operate with constant current, constant voltage, or constant conductance. For reasons that are explained in the sections on control circuitry, our proposed accelerometer operates at constant voltage in the low-bias millivolt regime.

II. Design Methodology

The process flow for the design of this accelerometer is shown in Figure 1 and can be described as follows. The requirement on full-scale acceleration and the chosen method of sensing dictate the required z-axis sensitivity of the device. This z-axis sensitivity then dictates the geometry of the proofmass. For a determined proofmass geometry, basic open-loop performance criteria can be evaluated, such as beam stiffnesses, modal frequencies, and cross-axis sensitivities. Once this proofmass geometry satisfies the basic performance criteria, more in-detailed design procedures and analyses are carried out to determine the electronics design, the pad layout, the tunneling tip geometry, the detailed microfabrication process flow, and characteristics associated with the detailed feedback simulation. This process is an iterative one, and much of the design involves tradeoffs between various parameters.

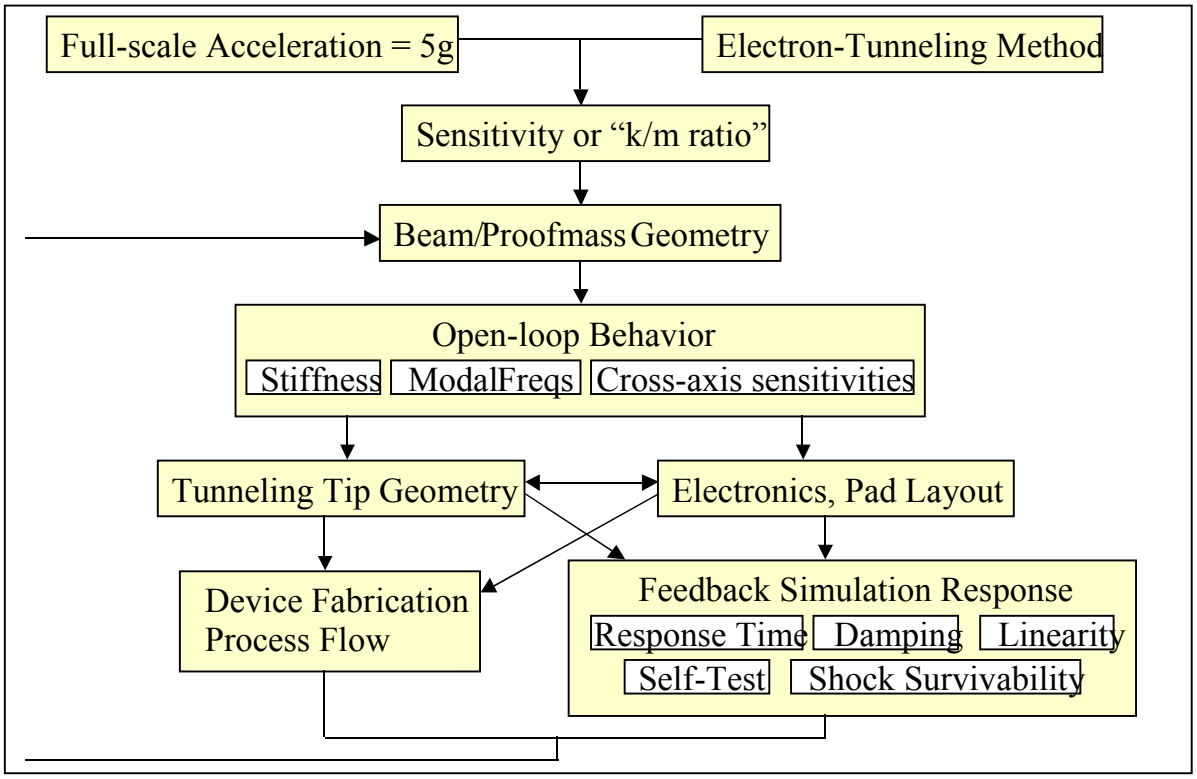


Figure 1: Design Process Flowchart

III. Proof Mass and Mechanical Design

III.1. General Discussion of Accelerometer Geometry and Specifications

The sensing means chosen for this proposed accelerometer is electron-tunneling current. This method of sensing acceleration requires the ability to control the motion of the device's proof mass within a displacement range of approximately 1 Angstrom. The bulk of successful accelerometers on the market today use other methods of sensing, such as capacitive or piezoresistive, which require significantly larger ranges of motion (on the order of hundreds or thousands of angstroms) than that required by electron-tunneling sensing. It is this displacement range that dictates the complexity of the proof mass design. As will be discussed in this section, whereas these other sensing types of accelerometers are required to possess large bulk micromachined proof masses suspended by intricately shaped tethers to successfully satisfy sensitivity specifications, the electron-tunneling accelerometer is able to use a much simper surface micromachined cantilever beam structure.

III.1.i. Accelerometer Sensitivity

An accelerometer's sensitivity is a function of the required magnitude of acceleration to be sensed and the method of sensing. The inverse of this sensitivity can be described by the "k/m ratio", where k is the stiffness of the proof mass spring and m is the mass of the proof mass. These characteristics are derived from a simple static force balance on the proof mass itself, where a is the magnitude of acceleration and z is the proofmass displacement:

$$F_{spring} = F_{acceleration}$$
$$kz = ma$$
$$\frac{k}{m} = \frac{a}{z}$$
$$"\frac{k}{m}\ ratio" = \frac{1}{sensitivity}$$

For example, given an accelerometer full-scale requirement of 5g and an electron-tunneling range of 1A, an estimate for the "k/m ratio" and sensitivity can be derived:

$$\frac{k}{m} = \frac{5g}{1A} = 5.0x10^{11}\ s^{-2}$$
$$or$$
$$sensitivity = 2.0x10^{-12}\ s^{2}$$

For capacitive or piezoresistive accelerometers under the same 5g requirement, since the displacement ranges are on the order of 1000Å, the required "k/m ratios" for these devices are 3 orders of magnitude smaller than for the proposed tunneling current accelerometer. As a result, these types of accelerometers require a combination of larger mass and reduced proof mass stiffness. Typical bulk micromachined proof mass geometries for capacitive or piezoresistive accelerometers are shown in Figure 2.

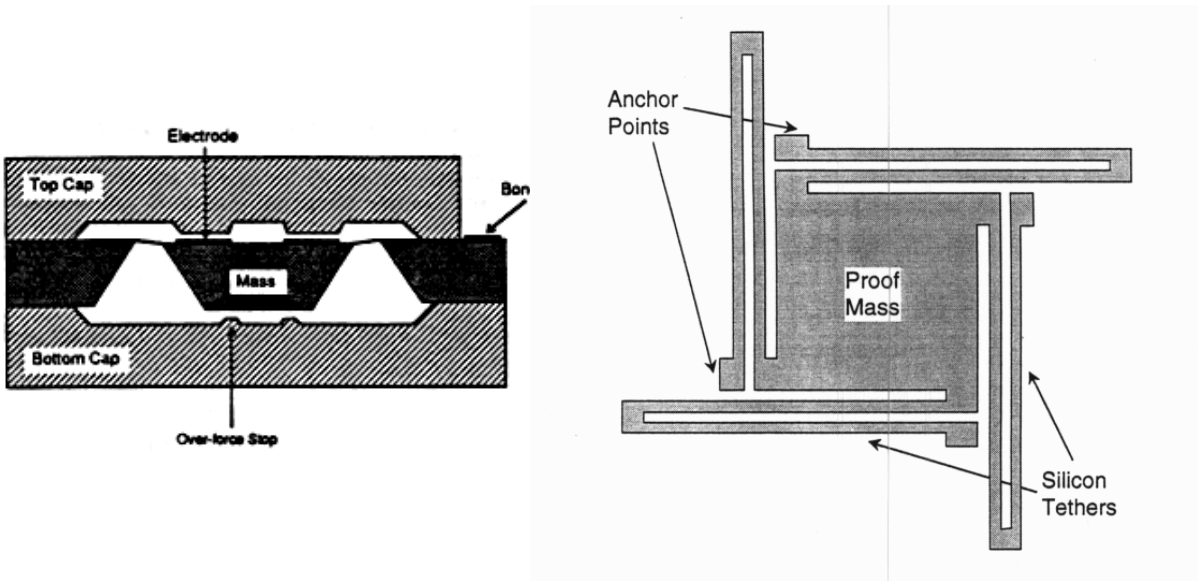


Figure 2: Typical Bulk-Micromachined Proof Mass Geometries (Left:Deb90; Right:Ash95)

These geometries are characterized by very large proof masses suspended by intricately shaped tethers. In comparison, the proposed electron tunneling accelerometer can use a simple surface micromachined cantilever beam to satisfy a much larger "k/m ratio" characterized by reduced mass and increased stiffness.

III.1.ii. Accelerometer Response Time

The response time of the device is dictated primarily by the natural frequency of the proof mass. In modeling a simple spring-mass system, the natural frequency is described by the following relation:

$$\omega_n = \sqrt{\frac{k}{m}}$$

As a result, accelerometers that possess very small sensitivities (or large "k/m ratios") inherently have high natural frequencies. For a simple spring mass second order system (assuming critical damping), the response time can be estimated by:

$$t_{response} = \frac{4}{\zeta \omega_n}$$

In reality, this expression is actually the definition of settling time, but for conservative estimates it is used here for the response time. Accelerometer's that possess high natural frequencies therefore exhibit very fast response times. Reorganizing this expression, and baselining a required response time specification , $t_{response}$, a lower bound for the required proof mass natural frequency can be derived:

$$\omega_n = \frac{4}{\zeta \, t_{response}}$$

Since the requirement for this accelerometer is $t_{response}$= 10 ms, the lower bound for required proof mass natural frequency is:

$$\omega_n = \frac{4}{(0.707)(10x10^3 s)} = 565\ r/s$$

This natural frequency puts the following requirement on the "k/m ratio" for the proof mass structure:

$$\frac{k}{m} \geq (565 r/s)^2$$

$$\frac{k}{m} \geq 3.2x10^5 s^{-2}$$

Based on these simple arguments, it is clear to conclude that if the proof mass structure satisfies the 5g acceleration requirement above, it will automatically satisfy the response time requirement. This is a benefit of using such a small displacement range technique as electron-tunneling current sensing.

III.1.iii. Cross-axis sensitivity

In any accelerometer, it is crucial that the z-axis sensitivity be at least two orders of magnitude smaller than any other axis. Otherwise, the desired z-axis output signal from the accelerometer will include unwanted cross-axis behavior. A significant advantage in using a simple constant thickness surface micromachined cantilever beam structure rather than a large bulk-micromachined proof mass design is the inherent reduction in this cross-axis excitation. The problem with most large bulk-micromachined proof mass structures is that the center of gravity of the proof mass is significantly far below the plane comprising the tethers that form the proof mass springs (again see Figure 2). Rocking motion is therefore a challenging issue to overcome. The cantilever beam design eliminates this problem altogether since the beam is of constant thickness throughout. The tunneling tip, located close to the beam free end, does protude below the beam, but because the tip is extremely small in volume compared to the beam dimensions, any transverse full-scale accelerations in the x or y directions produce minimal tip motion in the z-direction. Consequently, cross-axis behavior is minimized with such a cantilevered beam design.

Based on these concise arguments concerning z-axis sensitivity, response time, and cross-axis sensitivity, a cantilevered beam proof mass geometry is baselined.

III.2. Overview of Cantilevered Beam Design

The proposed electron-tunneling current accelerometer contains a cantilevered beam, shown in Figure 2, with the following dimensions, where L is length, W is width, t is thickness, gap is the spacing between the substrate and the underside of the beam, rho is material density, and E is the material Young's modulus:

L = 270 um
W = 60 um
t = 2 um
gap = 1.5 um
rho = 2300 kg/m3
E = 190e9 Pa

The tunneling tip is located at the centerline of the beam width and is spaced 10 um from the free end of the beam. The tip is 1.2 um in length and 2 um in diameter. An SOI approach was selected for creation of the beam due to the predictable properties and low built-in stress of single crystal Si (see Fabrication discussion). The base structure supporting the beam is comprised of SiO2 material and has dimensions 240 um x 240 um. x 1.5 um thick.

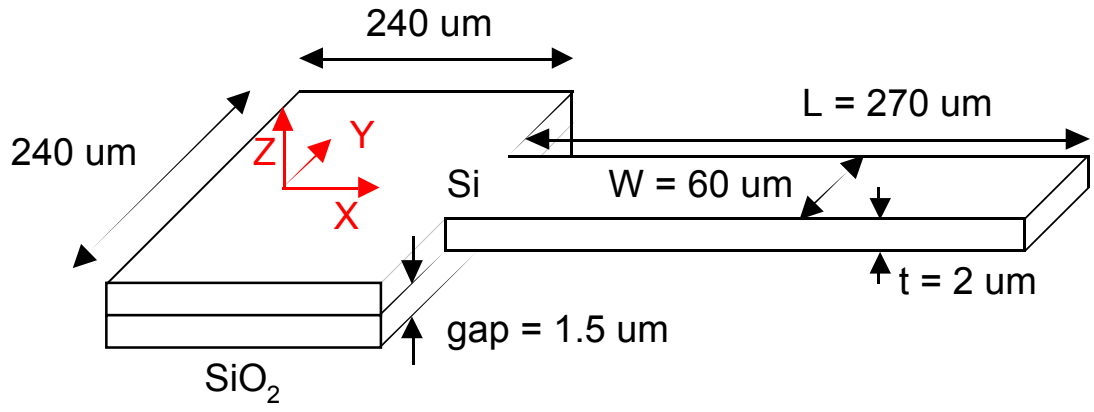


Figure 3: Final Proofmass Beam Geometry

Simple analytical models of the beam were initially used to estimate z-axis sensitivity, x-axis sensitivity, y-axis sensitivity, the cross-axis sensitivity and the beam modal frequencies. Using both analytical expressions and a detailed ANSYS finite-element model of the beam, tolerance analyses of the beam length, width, thickness, density, and Young's modulus were performed as well as investigations into the effect of the base structure dimensions, the base structure material, thickness variation along the beam, thickness variations across the beam, and tunneling tip placement. These analyses are included in Appendix A. Mode shapes of the final beam design are shown in Figure 3. The final characteristics of the cantilever beam structure (open-loop with no feedback) are now summarized.

- z-axis sensitivity = 2.412e-11 s^2 (or a "k/m ratio" = 4.145e10 s^{-2})
- y-axis sensitivity = 2.680e-14 s^2 (or a "k/m ratio" = 3.731e13 s^{-2})
- z-axis sensitivity = 3.577e-16 s^2 (or a "k/m ratio" = 2.796e15 s^{-2})
- Cross-axis sensitivity = 3.728e-17 s^2 (~0.0002% of z-axis sensitivity)
- Modal Frequencies (based on FEM): f1 = 40.748 kHz (1st z-axis bending)
  f2 = 258.53 kHz (2nd z-axis bending)
  f3 = 381.81 kHz (1st torsion)
  f4 = 765.40 kHz (3rd z-axis bending)

For this baseline final design, the 2nd mode frequency is 6.345 times the first mode frequency and therefore consideration of only the 1st mode dynamics in the full simulation is justified.

IV. System Model

There is a need to develop a finite-dimensional linear dynamic model of the system to be used in design analysis and feedback design. A more comprehensive model, including real-devices behavior, is required for better evaluation of the expected device performance. In what follows, models for the major system components will be presented.

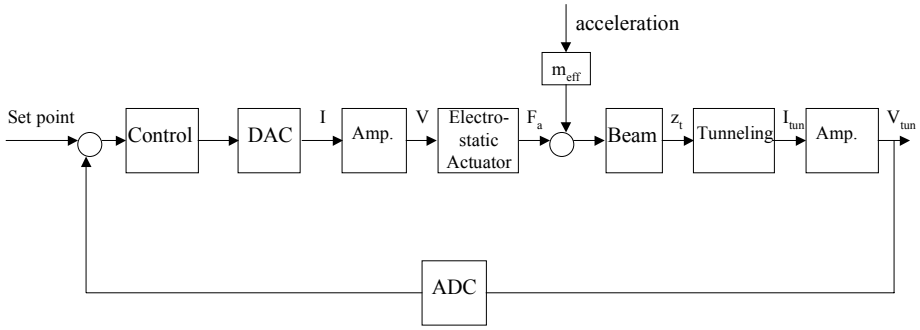


Figure 4: Block diagram of controller

IV.1. Beam Flexural Dynamics

The cantilever is modeled using elementary beam theory. In elementary beam theory, beams are assumed to have a tall cross-section, i.e. a plane stress condition. However, the beam under consideration is of a wide cross-section, i.e. a plane strain condition. The elementary beam theory is still applicable provided that the modulus of elasticity E, be replaced by $E/(1-\nu^2)$, where ν is Poisson's ratio. Consider the beam of Figure 5, with applied electrostatic actuation and self-test forces, F_a and F_{st}, respectively. The forces are modeled as concentrated loads acting at a distance x_a and x_{st}, respectively, measured from the fixed end of the beam. Damping force proportional to beam velocity relative to its base,

$$F_d = c\,\dot{z}(x,t),$$

is assumed. The effect of tip mass is neglected and justification will be given in (some where in Dave's section??). Using the principal of virtual work, and assuming a virtual displacement $\delta u(x,t)$, the governing equation of motion of the beam is given by,

$$\int_0^L \left(\rho A \frac{\partial^2 (z+z_b)}{\partial t^2}\delta u(x,t) + c\frac{\partial z}{\partial t}\delta u(x,t) + \frac{EI}{1-\nu^2}\frac{\partial^4 z}{\partial x^4}\frac{\partial^2(\delta u(x,t))}{\partial x^2}\right) dx = F_a\,\delta u(x_a,t) + F_{st}\,\delta u(x_{st},t)$$

where, z is the beam deflection relative to its base (fixed end), z_b is the motion of the base, ρ mass density, I is the moment of inertia, and $\delta(x-x_i)$ is the Dirac delta function which is infinity at $x = x_i$ and zero elsewhere.

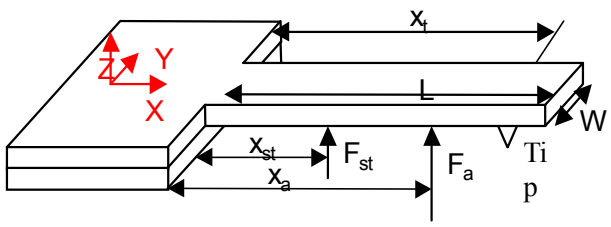


Figure 5: Schematic of the cantilever beam.

Using the method of assumed modes, the solution and the virtual displacement can both be expressed as,

$$z(x,t) = \sum_{i}^{\infty} \psi_i(x)\ q_i(t)\ , \qquad \delta u(x,t) = \sum_{i}^{\infty} \psi_i(x)\ \delta u_i(t)\,,$$

where, $\psi_i(x)$ is the mode shape function, and $q_i(t)$ is the temporal part of the solution. Substituting this form of the solution into the governing equation, and using orthogonality properties of the mode shape function, the resulting modal differential equation is,

$$m_i \ddot{q}_i + b_i \dot{q}_i + k_i q_i = \psi_i(x_a)\, F_a + \psi_i(x_{st})\, F_{st} - \rho\, A \int_0^L \psi_i(x)\, dx\ \ a$$

$$m_i = \rho\, A \int_0^L \psi_i^2(x)\, dx, \quad b_i = c \int_0^L \psi_i^2(x)\, dx, \quad k_i = \frac{E\, I}{1-\nu^2} \int_0^L \frac{\partial^2 \psi_i}{\partial x^2}\, dx, \quad a = \dot{z}_b$$

Using the boundary conditions of zero deflection and slope at the cantilever's fixed end, and zero moment and shear force at the free end, the mode shape function is found to be,

$$\psi_i(x) = (\sin(\lambda_i L) + \sinh(\lambda_i L))\,(\cos(\lambda_i x) - \cosh h(\lambda_i x)) \ + \\ (\cos(\lambda_i L) + \cosh(\lambda_i L))\,(\sinh(\lambda_i x) - \sin(\lambda_i x))$$

where, λ_i is the solution of the following transcendental equation, $\cos(\lambda_i L)\cosh(\lambda_i L) + 1 = 0$. The values for $\lambda_{1,2}$ L = 1.875, 4.694. The fundamental motion of the tip can be approximated using the first mode as,

$$z_t(t) = z(x_t, t) = \psi_1(x_t)\, q_1(t)$$

This assumption was later justified for the final beam design. The resulting transfer function from the applied forces to tip position is therefore,

$$Z_t(s) \frac{1}{m\, s^2 + b\, s + k} \left(\psi_1(x_t)\psi_1(x_a)F_a + \psi_1(x_t)\psi_1(x_{st})F_{st} - \rho\, A \psi_1(x_t) \int_0^L \psi_1(x) dx\ \ a \right) \qquad **$$

IV.2. Squeeze Film Damping

The dissipation in the mechanical structure is mainly due to squeeze film flow in the gap formed between the underside of the beam and the top of the substrate. In addition, squeeze film flow results in a spring-like force acting on the beam due to compressibility of fluid in the gap. If the bandwidth of the system is much smaller than a critical frequency ω_c, (i.e. low squeeze number σ), the spring force can be neglected compared to the dissipation force. In this case, the first mode damping coefficient in **, is given by,

$$b = \frac{96\mu L w^3}{\pi^4 g_o^{\ 3}}$$

$$\sigma = \frac{12\mu w^2}{P\ g_o^{\ 2}}\omega,\ \omega_c = \frac{\pi^2 g_o^{\ 2} P}{12\mu w^2}$$

where, μ is the fluid viscosity, w is the cantilever width, P is the nominal pressure, g_o is the nominal gap, and ω is the frequency. For the final beam design, ω_c = 460 kHz >> system bandwidth (~ 10 kHz), hence the spring force can be safely neglected.

IV.3. Electrostatic Actuators

The electrostatic forces are modeled as concentrated loads . To avoid pull-in effect, the tip length protruding below the beam is designed to be larger than two thirds of the nominal gap beneath the cantilever. The electrostatic forces are given by,

$$F_a = \frac{\varepsilon_o A_a V_a^2}{2(g_o + \frac{\psi_1(x_a)}{\psi_1(x_t)} z_t)^2}$$

$$F_{st} = \frac{\varepsilon_o A_{st} V_{st}^2}{2(g_o + \frac{\psi_1(x_a)}{\psi_1(x_t)} z_t)^2}$$

where, ε_o is the permittivity, $A_{a/st}$ is the area of the corresponding electrostatic pad, and $V_{a/st}$ is the voltage to the corresponding pad.

IV.4. Models for the feedback electronics

The tunneling current is sensed using a trans-impedance amplifier. The operational amplifier is modeled as single dominant pole with a gain-bandwidth product of $A_{o1}f_{o1}$. The effect of parasitic capacitance C_p, at the input of the amplifier is also modeled. The transfer function of this amplifier is given as,

$$\frac{V_{tun}}{I_{tun}} = \frac{-A_{o1}R_1}{R_1C_p\,\frac{s^2}{2\pi f_{o1}} + (\frac{1}{2\pi f_{o1}} + R_1C_p)\,s + (A_{o1}+1)}$$

The tunneling voltage is digitized through a DAC with a sampling frequency of f_{DAC}. The signal is processed and a control signal is then sent to an ADC with a sampling frequency of f_{ADC}. The main effect of sampling at a finite frequency is to introduce delays in the feedback loop. These delay will have little effect if the closed loop bandwidth is kept very small compared to the sampling frequency. Two sample-and-hold, (zero-order-hold), were used to model effects of ADC and DAC sampling. In addition, quantization effects due to finite bit-size were modeled as round-off quantizers. Moreover, manufacturer's data on least significant bit noise in the ADC was also included in the model random noise with an appropriate variance.

Output current from the ADC is sent to a second trans-impedance amplifier modeled with a transfer function of,

$$\frac{V_a}{I_{ADC}} = \frac{-A_{o2}R_2}{\frac{s}{2\pi f_{o2}} + (1+A_{o2})}$$

The voltage V_a is sent to the electrostatic actuation pad to compensate for changes in the tunneling current due to input acceleration.

IV.5. Beam Design: Dynamic Considerations

From the given design requirements, the following considerations are desired:

- High bending natural frequency.
- Well-damped bending mode for non-oscillatory transient response and small settling time.
- High torsional stiffness and natural frequency for low cross-axis sensitivity.
- Low electrostatic voltage requirement for DC deflection and self-test.

From the preliminary static analysis presented previously, the initial beam design had dimension of 225 x30x2 μm. The first modal frequencies are 58 kHz (Q ~ 6), and 450 kHz for the bending and torsional modes respectively. A potential process flow was developed for fabrication. This process required having holes in the cantilever for facilitate its release. Accordingly, the effective width for estimating squeeze film damping becomes the holes spacing (15 μm). This increases the quality factor Q ~ 48. In order to actively provide adequate damping (Q~1), for such a high-frequency mode a very high bandwidth electronics will be required. However, the gain-bandwidth product of the operational amplifiers and the parasitic capacitance impose a practical limit on the bandwidth of the electronics. A bandwidth of 400 kHz can be expected for the sensing trans-impedance amplifier for an estimated 1 pF parasitic capacitance. Therefore, it was necessary to change the proposed process to eliminate the release holes. This would allow increasing the width of the cantilever w, hence squeeze film damping ($\sim w^3$). The resulting trade-off is a decrease in torsional mode frequency ($\sim w^{-1}$), hence cross-axis sensitivity. In addition, the bending frequency was reduced by increasing the beam length. The final design has dimensions

that gave a good compromise between the aforementioned trade-offs. The dimensions are 270 x60x2 μm. The first modal frequencies are 40 kHz (Q ~ 1), and 381 kHz for the bending and torsional modes, respectively.

IV.6. Location of Electrostatic Pads

The location and size of the electrostatic pads were selected to provide a compromise between DC offset voltage and self-test voltage on one hand and avoiding interference of fringing fields between neighboring pads. The self-test pad is 30 μm long along the beam's length and has a 60 μm width and centered at 145 μm from the base of the beam. The actuation pad is 50 μm long along the beam's length and has a 60 μm width and centered at 205 μm from the base of the beam.

IV.7. Vacuum operation

Viscosity decreases at low pressures. At atmospheric pressure the viscosity of air is 1.8×10^{-5}. At 300 ^{o}K and 10 Torr, the viscosity of air decreases to 1.6×10^{-6}. From dynamic response considerations, operation in vacuum is very difficult to achieve due the limited bandwidth of the feedback electronics.

Nitrogen and air have very similar values of viscosity at room temperature ($\mu_{air} = 1.84 \times 10^{-5}$, $\mu_{air} = 1.77 \times 10^{-5}$). Therefore the model for the device is nearly identical for operation in nitrogen or air environments.

IV.8. Linearization and tunneling set point:

It is a common practice in STM machines and in some tunneling accelerometers (Yeh et. al.) to use a logarithmic amplifier to linearize the exponential dependence of the tunneling current on the tunneling gap. This approach was not adopted here, since it would introduce additional dynamics that may reduce the maximum achievable bandwidth. Alternatively, a high-bandwidth linear controller with a high gain-margin will be designed to compensate for this nonlinear tunneling behavior. The tunneling current set point was selected as compromise between good device sensitivity and a large gap size to avoid tip crash during transients.

IV.9. Controller Design

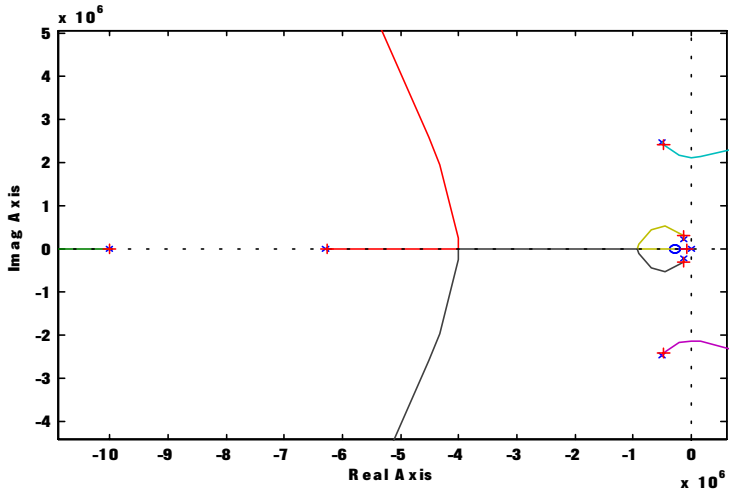
The controller is expected to maintain the tunneling current constant in the presence of input acceleration. It should provide as fast response as possible with a well-damped transient to avoid tip crashing and long settling time. The control problem can be thought of as disturbance rejection problem, with the input acceleration acting as a disturbance. To be able to reject this disturbance the closed loop bandwidth has to be large compared to the frequency of the measured acceleration. In addition, to provide active damping of the cantilever dynamics, one can canceling the effect of these dynamics by introducing controller zeros with a frequency that is very close to the cantilever's modal frequency. This approach of stable pole-zero cancellation is however, not practically robust. Uncertainties in the location of the cantilever may result in high frequency ringing. Alternatively, by placing controller zeros close to the real axis, lightly

damped modes can be adequately damped. This typically requires a large bandwidth. However, there is a practical upper bound on the closed loop bandwidth. The bandwidth is limited to be considerably lower than feedback amplifiers, bending and torsional resonances of the beam.

The structure of the controller is a PID with an additional high-frequency pole. The integrator is used to reduce sensitivity to op-amps input voltage drift and offset and other constant or slowly varying disturbances. The zeros of the controller provide the required phase-lead near crossover frequency. The sampling frequency (1 MHz) is very large compared to the expected bandwidth (few kHz). Therefore, the controller can be designed in continuous time and then digitized to obtain the equivalent discrete-time controller with no degradation in closed loop performance.

During transients, large changes in the effective loop gain are expected due to the exponential dependence of tunneling current on tip displacement. For example, for a tunneling gap of 8.5 ^{o}A, $I_{tun} \approx 12\ \Delta_{zt}$, where at a gap of 5 ^{o}A, $I_{tun} \approx 52\ \Delta z_t$. Moreover, the tunneling barrier height ϕ, is a very uncertain parameter. For a gold tip tunneling in air, changes as much as a factor of 5 has been reported (Ken). Hence, it is required to be robust to gain uncertainties by designing for a large gain margin. Additionally, the effect of sampling and higher order dynamics require that the bandwidth be limited and to have a large phase margin.

A system model linearized at the operating point was used for controller design. The root locus plot is shown in Figure 3. The linear frequency response is shown in Figure 4. The closed loop system has a gain margin of 8, a phase margin of 92^{o} and a crossover frequency of 9 kHz.



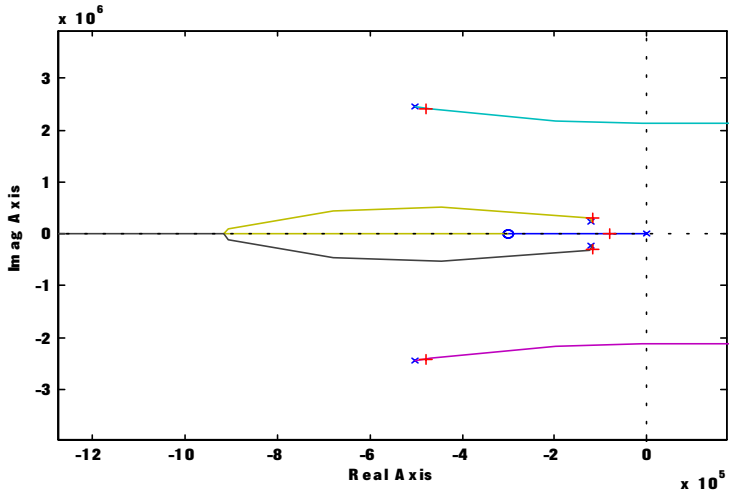


Figure 6: (top) Root locus plot, (bottom) a zoom-in view.

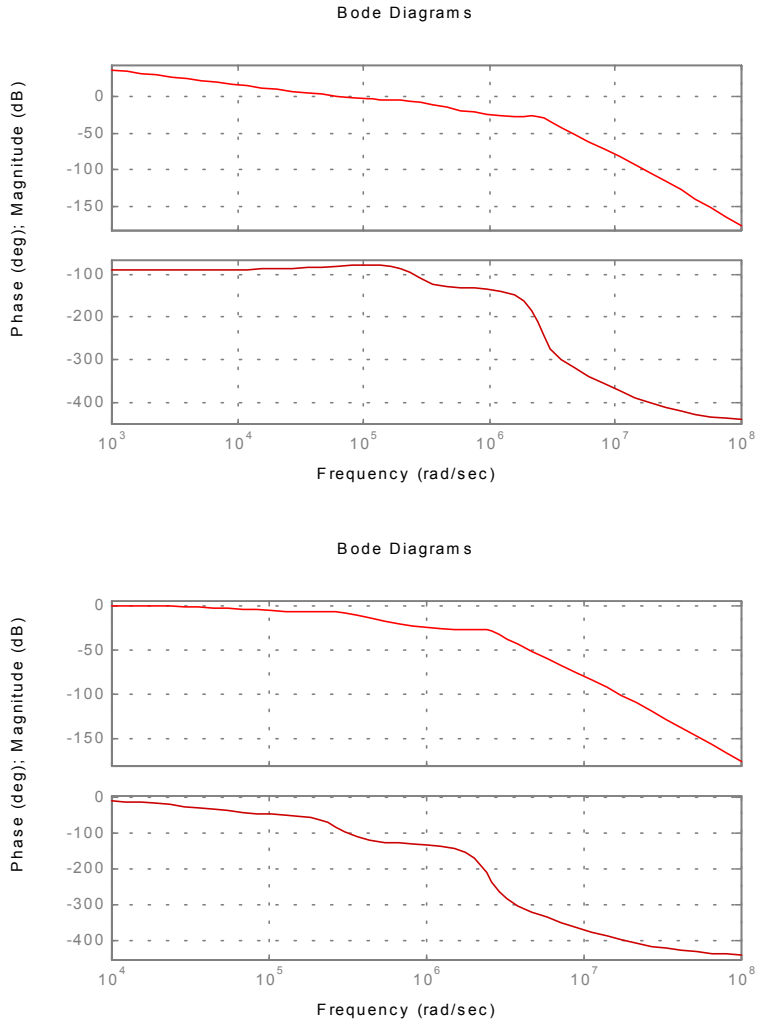


Figure 7: Linear frequency response, (top) open loop, (top) closed loop from set point to tunneling voltage.

IV.10. Simulation Results

The full nonlinear model, including noise and sampling effects was used to simulate the expected device performance. The main observations are presented

Step response: ± 5g, **Figure 8.**

- Amplifier offset voltage shifts the operating point from 8.5 Å to 8.86 Å.
- Initial response is different for positive or negative acceleration signals due to changes in effective loop gain during transients.
- Steady state response is repeatable and consistent.
- Response time is 35 μs for positive and negative acceleration signals.

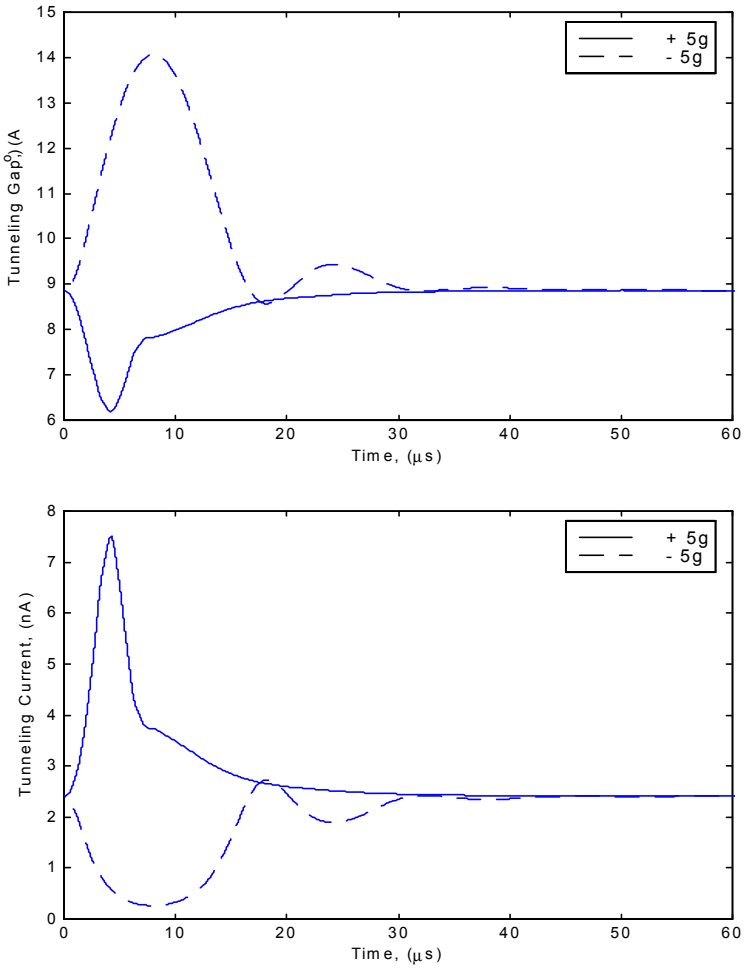


Figure 8:Step response for 5g acceleration

Full-scale nonlinear response: (**Figure 9**)

- Without any changes to the nominal design, the device can used in low acceleration high bandwidth or high acceleration low bandwidth applications, e.g. 5g upto 10 kHz, or 50g upto 1 kHz.
- Transients dies out quickly.
- Device has excellent linearity (control voltage vs. input acceleration).
-

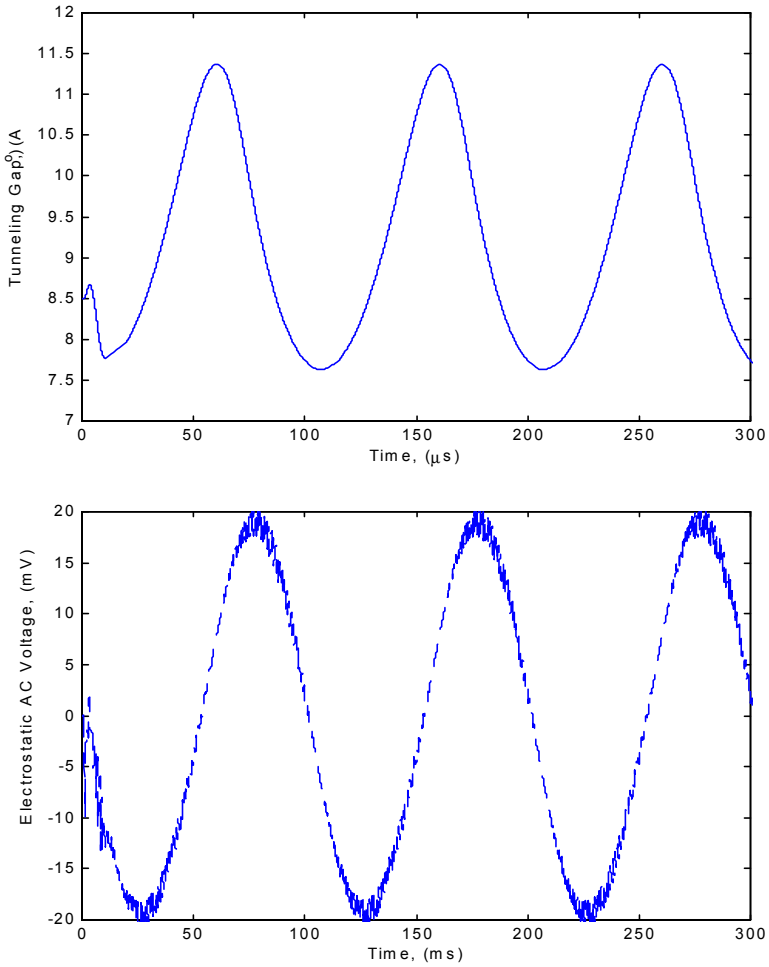


Figure 9:Nonlinear frequency response, 5g 10 kHz.

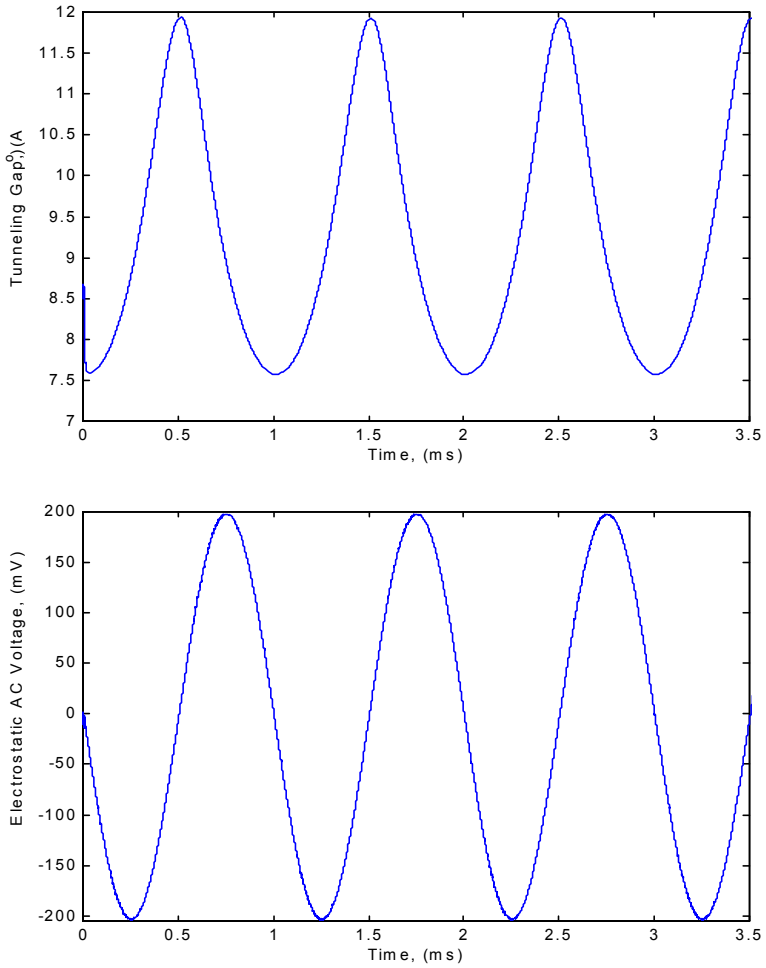


Figure 10:Nonlinear frequency response, 50g 1 kHz.

Minimum detectable Acceleration: 20 mg (Figure 6)

- Limited mainly by LSB and amplifiers noise (~ 20 μV).
- Minimum detectable acceleration 20mg, (0.1 % for 5g full scale or 0.01% for 50g full scale).

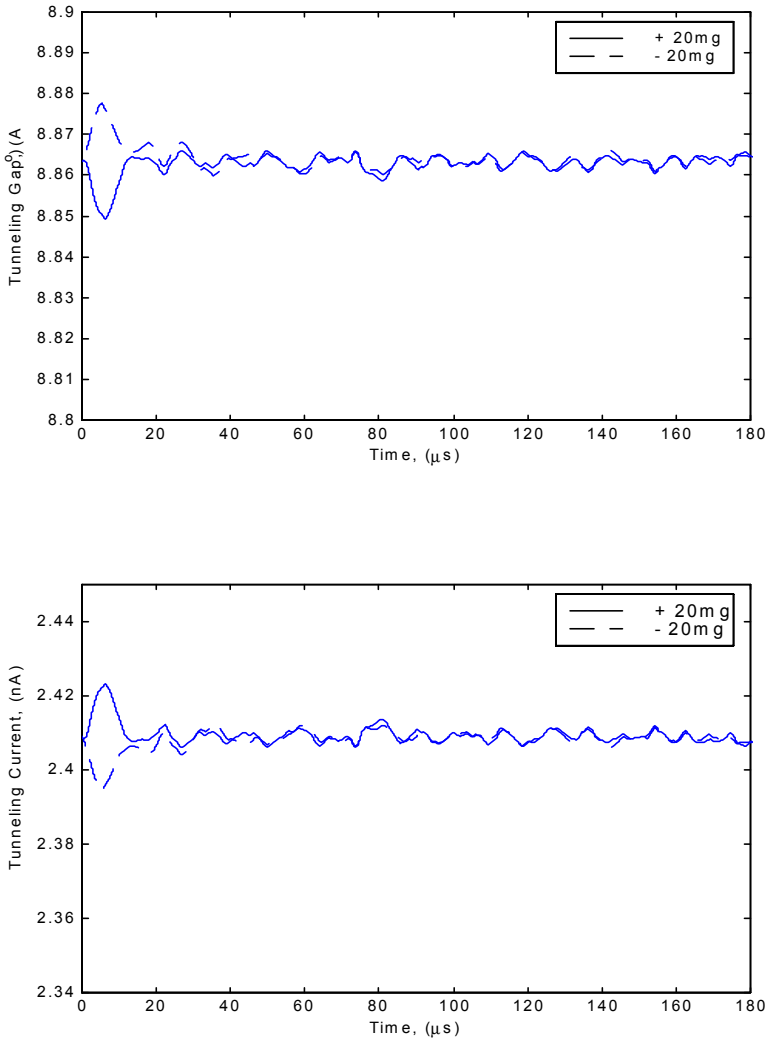


Figure 11:Step response for 20mg acceleration

Shock test: 100g held for 100 ms with 1 ms rise time (**Figure 12**)

- Very small deviation from nominal position during ramping (< 0.5 Å) .
- Controller maintains the gap at nominal value during hold period, therefore, simulation is shown for 5 ms hold period instead of 100 ms.

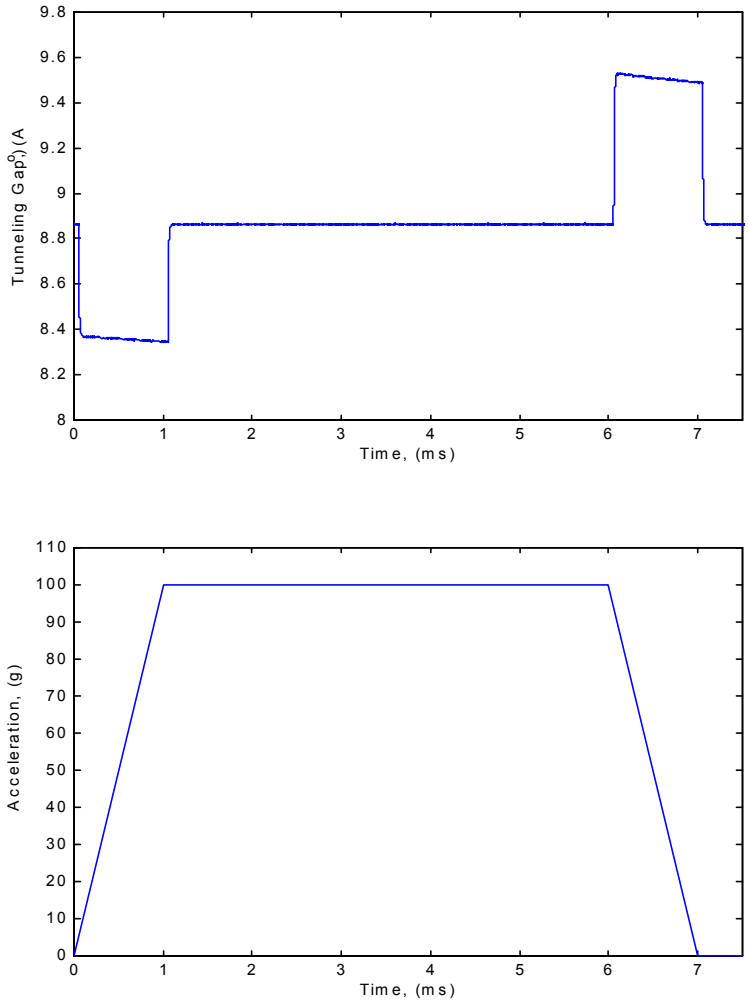


Figure 12:Shock test, 100g, 5 ms, 1 ms rise time.

Robustness to parameter uncertainty:
Variations in the tunneling barrier height ϕ: 5g step response, ϕ = 0.5 eV (**Figure 13**), ϕ = (**Figure 14).**

- Variations in the tunneling barrier height for a gold tip in air between 0.05-0.5 eV, form a nominal value of 0.17 Ev, (Ken).
- Excellent feedback stability robustness to large perturbations, (-70 % to +300 %).
- Moderate changes in the transient behavior with consistent and steady state response.
- Slight increase in response time from 35 to 50 μs for a –70% variation.
- Changes in the nominal gap from 8.5 Å to 5 Å (0.5 eV), and 16 Å (0.05 eV).
- Maximum expected value of ϕ has to be considered in selecting nominal operating point, to avoid tip crash during transient.

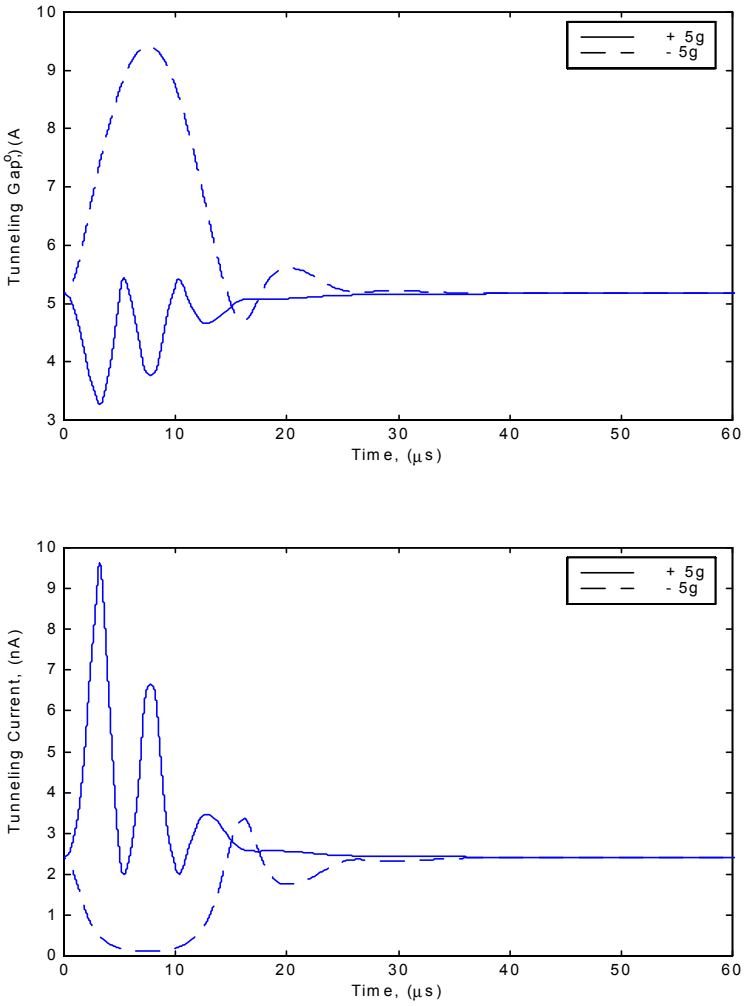
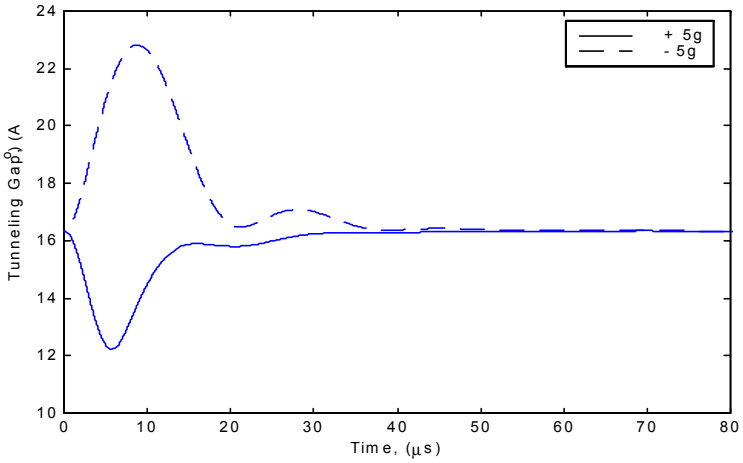


Figure 13:Step response for 5g acceleration, with variation in tunneling barrier height, 0.5 eV



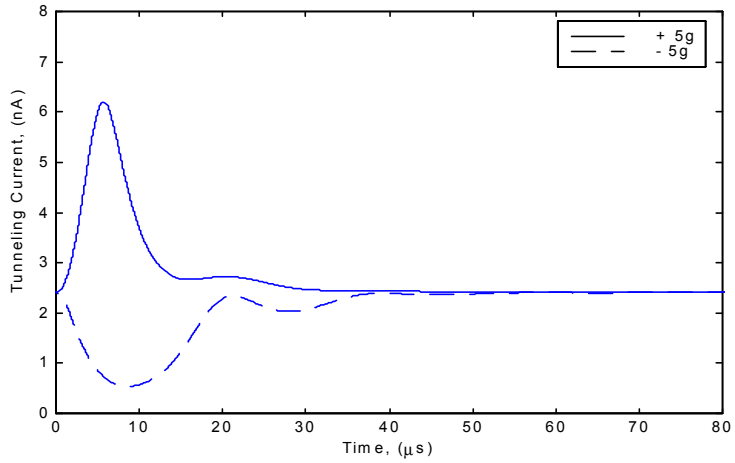


Figure 14:Step response for 5g acceleration, with -70 % variation in tunneling barrier height (0.05 eV).

Variations in squeeze film damping: 50% reduction in damping coefficient (**Figure 15**).

- Excellent feedback stability robustness to large perturbations, despite a 50% reduction in damping coefficient.
- Moderate changes in the transient behavior with consistent and steady state response.
- Slight increase in response time from 35 to 55 μs.

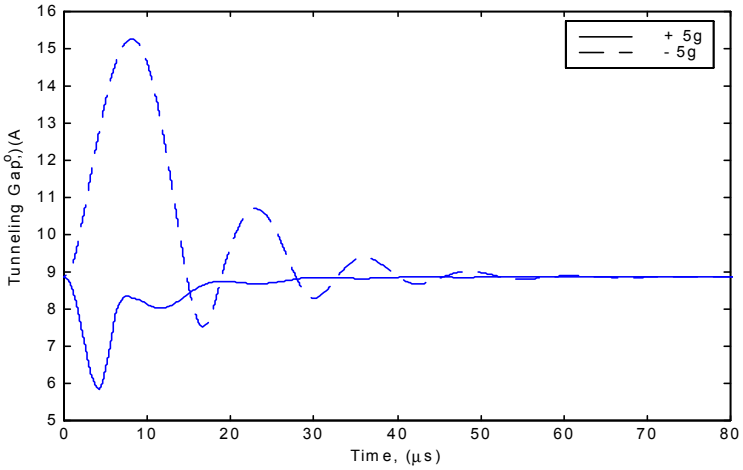


Figure 15:Step response for 5g acceleration, with 50 % decrease in squeeze film damping.

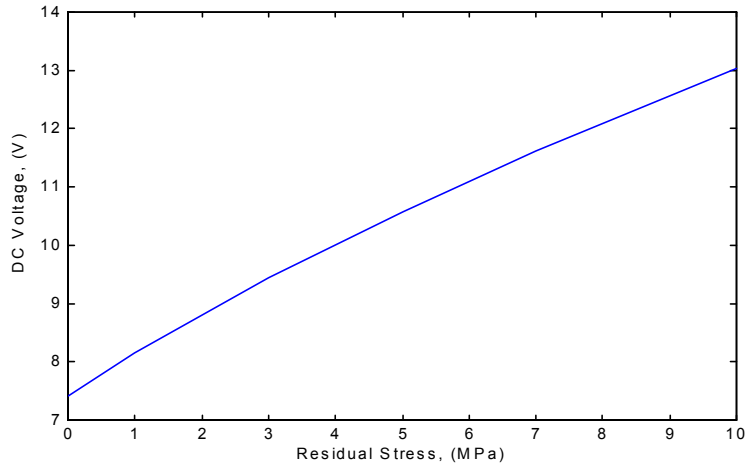


Figure 16: Residual stress, if present

Device sensitivity and linearity: (**Figure 17**)

- Sensitivity of 4 mV/g (noise ~ 20 μV, min. acceleration 20 mg 80 μV).
- Lineraity << 1%, is easilly achieved. Final device linearity may depend on signal conditioning.

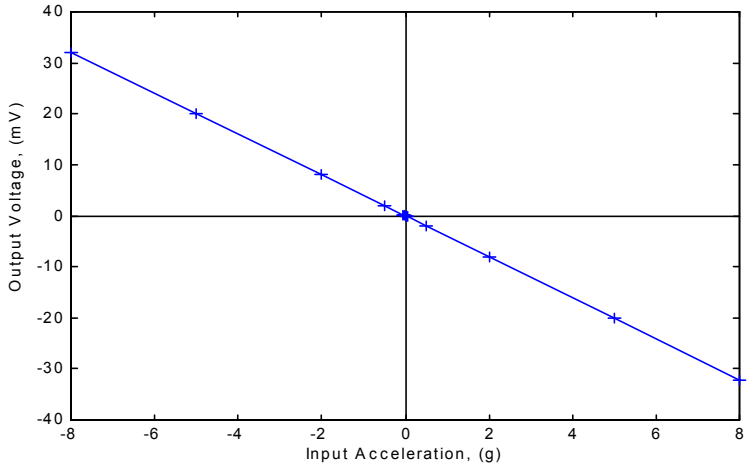


Figure 17:Input acceleration vs. output voltage.

V. Electronics

V.1. Design Principles

Because of the wide variety of parameters upon which the high-sensitivity tunneling current depends, including the CVD oxide thickness (and hence the actuation capacitor dimensions), the length of the tip (which depends on a timed oxide etch), the beam dimensions and material parameters, and the history of the tip (a crash, albeit unlikely given the soundness of the control algorithm, might require recalibration), we have chosen a digital core for the signal chain.

This permits us to set DC voltages such as the large component of an electrostatic actuation voltage, for example, with a certain number of bits in an EEPROM, as opposed to laser-trimming a resistor in a voltage divider, or requiring the end-user to compensate with external circuitry. It also allows us to change the behavior of the device for different regimes of operation – as noted earlier, we can operate the accelerometer in a high-bandwidth, low-acceleration mode, or a low-bandwidth, shock-acceleration mode. DACs and ADCs are fairly cheap, and offer a large degree of control for very little investment.

There are four major voltages that need to be specified: the first is the voltage of the beam (and thus the tip), V_{beam}, which is directly fed from the power supply. The voltage of the self-test pad, V_{self}, is specified by an EEPROM value which is fed into an 8-bit DAC, which specifies voltages in the range of 0.63 to 0.77 V (assuming a possible 10% error about the exact value of 0.70 V for a 2.5-g force), in increments of 0.55 mV; this permits us to specify the value of the self-test acceleration to better than 0.08 %. The actuator pad, V_{act}, has a DC offset of 10 V, which brings the 1.3 μm-long tip down roughly 200 nm, so that it is only a few Angstroms away; this V_{act} is precisely found by ramping up the voltage until a tunneling current of 1 nA appears; the resultant value can be stored in the EEPROM, to be fed out through an 8-bit DAC. For a possible error of 20%, we would specify voltages in a range roughly spanning 8 V to 12 V, with a resolution of 15 mV. (The small control voltage which is added to the DC component of the actuator voltage, operates roughly between 15 mV and –15 mV in normal circumstances; we control this with a high-speed 14-bit DAC which is operated so that the most significant bit of the small control voltage is greater than the least significant bit of the DC component of the actuator voltage.) Finally, the tunneling-pad voltage V_{pad} is specified either by a linearly voltage-divided Zener diode, or by some other relatively constant source, with respect to V_{beam}. As we will see in the section on instrumentation, the tunneling-pad voltage is actually specified as a virtual voltage through a transimpedance amplifier, so that the tip-to-pad voltage is held at around 0.2 V.

V.2. Single-conductor Design

One key aspect of the electronics design is that we use a single conductor for both the tunneling tip and the part of the beam which responds to the capacitive actuators. This is only possible for a well-decoupled system – that is, one for which the various current paths do not interfere. Since we are operating at low frequencies, at first glance this seems like it shouldn't be a problem, but it is worth to consider possible effects of using a single-conductor design, since our signal and power currents are phenomenally low. As we will see, despite the low currents, the resistances and capacitances are of appropriate scales that we need not worry about couplings; the beam can be treated much as ground planes are in radiofrequency design.

Why use a single conductor? Creating separate conductors for different current paths would add another mask, complicate the beam model, and possibly exacerbate registration issues since we would be trying to fit three separate current paths on a beam that measures a few dozen microns wide. The most important reason, however, is that it's not necessary, as we'll see from the following brief arguments.

The capacitance between the R = 1 μm-radius round tunneling tip (as specified in the process flow) and a tunneling pad d_0 = 10 Å below it, is

$$\varepsilon_0 \int \frac{dA}{d} = \varepsilon_0 \int_0^{2\pi} d\theta \int_0^R \frac{rdr}{\sqrt{R^2 - r^2} + d_0} = 0.26 \text{ fF}$$

whereas the capacitance between the actuation pad and the beam is

$$\frac{\varepsilon_0 A}{d} = (8.854\times10^{-12})(60\times10^{-6})(30\times10^{-6})/(1.5\times10^{-6}) = 11 \text{ fF}$$

which is significantly more. In any case, both of these values present an incredibly high impedance at the relevant frequencies of a few kilohertz ($1/Cs$ ~ 50 teraohm), and thus cross-coupling can be neglected, since the resultant currents would be much smaller than the currents of interest anyway. Further, the tunneling current, which obeys the proportionality

$$I_{tun} = g_0 V e^{-\alpha\sqrt{\phi}x}$$

gives a dynamic resistance at 10 Å of over 200 MΩ, whereas the output resistance of the power supplies and opamps is a few ohms at most. Finally, we can estimate the relative values of the relevant currents, using results from the simulation, which suggests that in a 5-g acceleration, the control loop changes the voltage of the bottom actuator pad by 13 mV over 0.15 ms:

$$I = C_{act} \frac{dV_{act}}{dt} = (10.6\,fF)\frac{13mV}{.15ms} = 0.92\,pA << I_{tun} \sim 1nA$$

and so we see that in practice, the capacitive current would not affect the tunneling current signal by more than 0.09%, even if the current paths were not well-decoupled.

V.3. Instrumentation: elements

Although we would almost certainly choose to create an ASIC for this MEMS device, and operate the device in a two-chip configuration, for the purposes of modeling we decided to use commercially available parts with existant specifications. The reasons are twofold: first of all, we need fairly sensitive components in order to acquire and synthesize signals on the tiny scales of nanoampere tunneling currents and angstrom-size displacement, and although the technology for making these very sensitive parts exists, to meet all of their specifications might take decades of iteration and sharing of insight. Secondly, if we were to replicate these parts, the specifications

would be very similar to the original parts (with a few exceptions, like die size and power consumption, since we don't need the generality of an end-user IC), given that we are operating near some of the fundamental physical limits for room-temperature devices, so for modeling purposes there is nothing to be lost by using existant parts with well-defined models.

We have chosen four commercially available parts with published macromodels, for the signal chain of our system. The 1 nA-scale current first flows through a transimpedance amplifier consisting of a 1 MΩ resistor (Johnson noise of 0.129 μV/Hz$^{1/2}$) which connects the - terminal of an opamp to its output, with the + input at V_{pad}. The negative terminal, which is connected to the tunneling pad, is therefore held at the virtual voltage V_{pad}. We assume a parasitic capacitance of 1 pF over this contact, which we include in the control model. For the macromodel for this amplifier, we use that of the AD515 opamp (Ana), which draws a mere 75 fA of input current due to its extremely high-impedance JFET front end. The AD515 has a 1 MHz gain bandwidth product, which we compensate so as to give an open-loop gain of 1000 and a dominant pole at 1 kHz. The noise power spectral density is 55 nV/Hz$^{1/2}$ across 10 Hz-1 kHz, and the opamp has a 0.4 mV offset voltage, which we digitally subtract off in the digitization process.

The ADC model that we are using is that of the AD9260 (Ana), which can sample at 2.5 MHz with 16-bit resolution. It has a sigma-delta architecture with a flash quantizer, a pipelined A/D with 5 successive approximation stages, and is on the whole very complex. However, this type of A/D converter is now becoming fairly common, spurred onward by the telecommunications revolution, and it is likely that in the coming years such digitzation capabilities will be common in many devices. In particular, this model has about 30.5 uV/Hz$^{1/2}$ of input noise, for a 2 V full scale signal, which is a useful and reasonable number to incorporate into our model.

The DAC model for our sophisticated 14-bit DAC is that of the AD9754 (Ana), a 14-bit, 125 MHz DAC, which can source $\pm$ 20 mA of current, with 50 pA/Hz$^{1/2}$ current noise. As before, such DACs are now becoming quite accessible, and will make digital control more accessible for high-sensitivity, high-bandwidth applications. The fast conversion rates of the ADC and DAC allow minimal effects from the zero-order hold effects of analog/digital interface circuitry, thus allowing us to preserve fast settling times on the order of tens of microseconds.

The DAC output is fed into a 10 ohm resistor (with negligible Johnson noise, 0.4 nV/Hz$^{1/2}$), for a final transimpedance amplifier stage which creates the control voltage. The + terminal of the transimpedance stage is held at the actuator DC offset voltage (~10 V, for bringing the beam down 200 nm, as observed in the previous section), so that the output of the transimpedance stage provides 10 V $\pm$ 200 mV, which is sufficient to accurately servo all accelerations on the order of 5 g ($\pm$ 15 mV), while still having enough range to compensate for a 100g shock. The amplifier model we used for this stage was the AD743, a low-noise opamp which has only an input noise of 3.5 nV/Hz$^{1/2}$, while still preserving a gain bandwidth product of greater than 1 MHz (which we compensate, again, as having a gain of 1000 with a dominant pole at 1 kHz).

Also important for the system is a switching power supply, to provide step-up conversion of the input power to whatever voltages are required for the actuation, and possibly a charge pump, if extremely high voltages are needed – although from the above analysis, and from the simulations, it appears that this requirement is superfluous.

Schematics and explanations of DACs, ADCs, high-impedance and low-noise opamps, Johnson noise, switching power supplies, and capacitor-diode charge pumps can all be found in (Hor89).

V.4. Instrumentation: model

On the next page, we now present a schematic for the circuitry, and explain the components as necessary. The model was implemented in Simulink, as explained earlier. All of the noise values and bandwidth parameters in the previous section, as well as the effects of zero-order-holds, and the parasitic capacitance were simulated.

The p++ beam has resistivity ~ 10^{-3} Ω-cm, which corresponds to a 9 Ω resistance between the base of the beam and the point on the beam opposite the self-test pad, a 10.4 Ω resistance between the self-test pad and the actuator pad, and a 2.1 Ω resistance between the actuator pad and the gold tip. Since the current flowing through the beam is order 1 nA at all times, the voltage drop across the beam is 31.5 nV, which is far below the noise floor due to the amplifiers and resistors, and much less than the actuation voltages, and so we can completely neglect the resistance of the beam, and treat the system as well-decoupled.

A 1 pF parasitic capacitance was chosen as a reasonable value, given sources in the literature. The transimpedance amplifier (AD515) converts the 1 nA current into a ~ 1 mV signal, which is then digitized by the ADC at 1 MHz. A digital PID controller (whose implementation is described in the next section) operates on this signal, which is then output at 10 MHz through a DAC. The increased speed of the output was used to reduce the effects of the zero-order-hold on the output; a simple and perhaps wise alternative might be to put a low-pass filter on the output of the DAC, at a reasonably low frequency (say, 50 kHz), but this was not modeled, and didn't prove necessary given the very low frequency of the mechanical system compared to the capabilities of modern digital/analog interface circuitry. The DAC output was then inserted into a transimpedance stage/summer amplifier, which created the voltage for the actuation pad.

In an EEPROM are stored bits indicating the value of the self-test voltage, the DC offset voltage, the various switching power supply timing parameters, the exact values of the parameters used in the PID controller, the controller setpoint, and the values of the various offset voltages and currents that need to be added or subtracted from the default values in order to run the accelerometer at maximum performance. We discuss how to set these values in the next section.

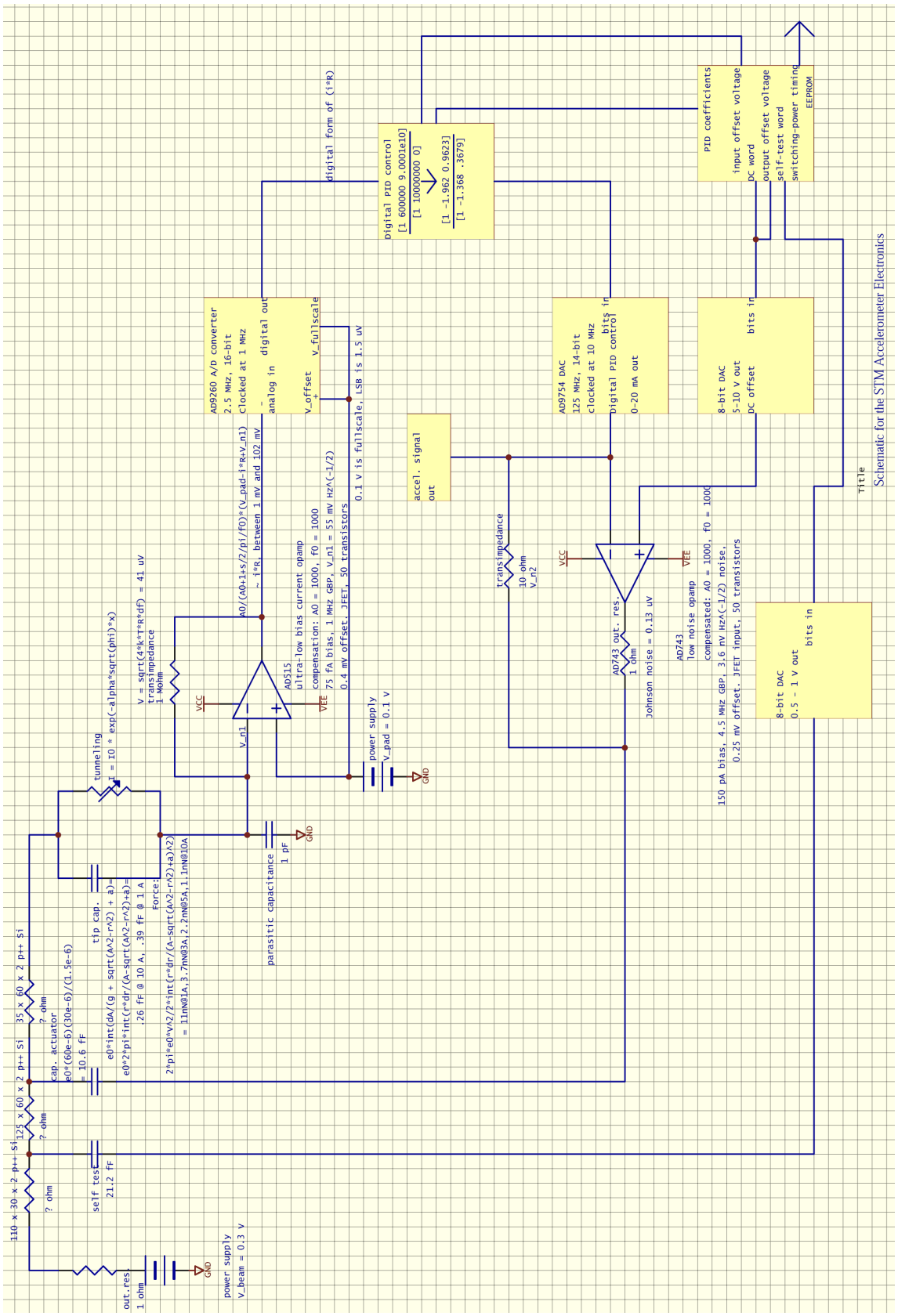


Figure 18: The Electronics

V.5. Instrumentation: implementation

We very briefly speculate on the size of the resultant package, since we were asked to keep the mass of the system under 5 grams. We also noted that each of these devices can be made using a single layer of circuitry, and therefore is suitable for SMARTMUMPS style fabrication processes. Each of the two opamps requires on the order of 50 transistors, and can be expected to take up an area of 1 mm^2 each. Each DAC requires on the order of 30 resistors, plus 15 transistors to act as current sources, plus 30 transistors to make a simple summing amplifier, and some overhead circuitry; this can reasonably fit in an area of about 1-2 mm^2. The ADC requires several dozen transistors and resistors, and might take up an area of 2-3 mm^2. Finally, the digital logic as synthesized in Verilog takes up around 400 gates, when explicitly optimized so that that as much as possible, bits are routed so that no computation is necessary. For example, if our digital filter is

$$I_o(n) = V_i(n) - 1.962\ V_i(n-1) + 0.9623\ V_i(n-2) + 1.368\ I_o(n-1) - 0.3679\ I_o(n-2),$$

which in fact it is (all gain parameters are external, in the amplifiers), then we can implement, for example, multiplication by 1.962 as multiplication by 1.28, plus multiplication by 0.64, plus multiplication by 0.004, and so on. Each multiplication by a power of 2 corresponds simply to a bit-shift operation, which can be hardwired (so that it doesn't require any gates) to the input of a simple adder. A more versatile programmable version would require more gates, but would still probably fit into a square 2 mm on a side if the feature size were 1.2 μm.

The total area is therefore just under 3.5 mm by 3.5 mm on a side. We may have been overoptimistic on some of these estimations, but we are probably not far off. And certainly we are far under the 5 g limit.

We also briefly explain the test-and-calibration method. First, apply a 0.2 V voltage between the tunneling tip and pad and increase V_{act} until tunneling appears, at the 1 nA level. That is the setpoint for the 8-bit DAC that synthesizes the DC-offset component of the actuator pad. Measure the offset voltages of the amplifiers for baseline subtraction. Then if necessary (although it is unlikely), tune the controller to insure a proper settling time, final value, and transient signal. Laser trim the resistors to provide appropriate transimpedance gains. Calibrate the readout by adjusting the acceleration setting on a shaker table and reading out the voltage reported by the accelerometer electronics – since after the settling time passes,

$$F_{net} = k\Delta x = 0,\ ma = F_{act},$$

and the digitally synthesized actuator force directly reflects the acceleration, so that a simple linear fit between actuator force and input acceleration should suffice to calibrate the part. Finally, measure the voltage that must be applied to the self-test pad in order to cause a 2.5-g acceleration, as compared to a standard 2.5 g acceleration created by the shaker table.

VI. Process Flow

VI. 1. Overview

We have developed a five mask process flow for the fabrication of the microelectromechanical assembly for a tunneling accelerometer. A cantilever proof mass with an integrated tunneling tip is created within a single bonded wafer. The dimensions of the completed proof mass are 270 × 60 × 2 μm with a nominal 1.5 μm gap between the actuation, self-test, and sense electrodes and the underside of a conductive p^{++} epi-Si cantilever. The tunneling tip protrudes about 1.3 μm from the underside of the cantilever.

The process flow described in this section satisfies three major challenges. The first challenge was set forth in the initial design problem statement. Specifically, the tunneling tip and the proof mass must are fabricated from the same wafer. The squeeze film damping requirements of accelerometer dictated the second challenge. Etch release holes in the beam drastically reduce the squeeze film damping effects. Therefore, the cantilever is released without the aide of etch release holes. The third challenge was self-imposed by the design team. Specifically, every effort has been made to design a simple process flow that uses the fewest number of masks or process steps, and does not require unusual or expensive fabrication technologies. The expected benefits of meeting this third challenge include less uncertainty in the fabrication process, improved tolerance benchmarks, reduced time to market, and lower cost per die, among others.

A conventional doped SOI wafer can be fabricated by bonding a first silicon wafer with an oxide layer to a second silicon wafer with a p^{++} epi-Si layer so that a bond line is formed along an oxide-epi interface. We have implemented a process flow that is based on a modified SOI wafer fabrication procedure. Before the bond is made, four process steps are carried out. First, a nitride dielectric layer is deposited. Second, gold metal (with two additional adhesion promoting metals) is deposited and patterned to form the electrodes and contact pads. Third, an oxide layer is deposited. Fourth, a cavity is etched into the oxide layer just below the point where the cantilever is to be formed. This presence of this cavity significantly reduces the lateral dimension of oxide that must be removed during the final cantilever release process. Therefore, etch release holes are not required and the second challenge is satisfied. Finally, this wafer is planarized and polished and then bonded to a second commercially prepared wafer with a 2.0 μm p^{++} epi-Si layer on one surface. The two wafers are fused together along an oxide-epi bond line without the use of special alignment procedures. This modified SOI wafer is then thinned down to the p^{++} epi-Si layer with a combination of grinding and anisotropic etching. The tunneling tip is formed by first etching a hole through the p^{++} epi-Si layer. The hole then acts as an etch mask so that a spherical depression can be isotropically etched into the underlying oxide layer under high agitation. The depression is used as mold for subsequent deposition of gold. This forms a gold tunneling tip with a spherical end. Then the lateral dimensions of the cantilever are defined. The wafer is bonded to wafer mounting tape and the die is cut using a die saw. A timed oxide etch is used to release the cantilever and to expose the end of the tunneling tip protruding from the underside of the cantilever. This is followed by a super critical CO_2 release step. Finally, the cantilever chip and the electronics chip are mounted side-by-side in a two chip package, ball bonded, and sealed under nitrogen.

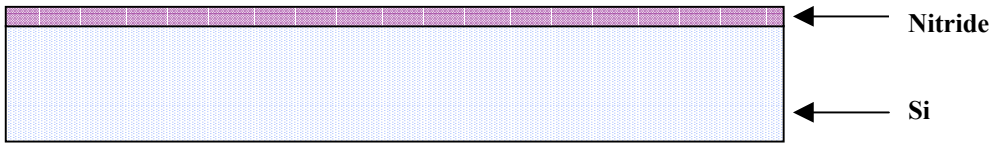
We have also designed an alternate process flow for the fabrication of a tunneling accelerometer with a tunneling tip that has more precise geometry and smaller tip radius. This process flow is similar to the process flow presented here. However, the primary modification is that the mold for the tunneling tip is anisotropically etched in to an undoped epi-silicon layer instead of isotropically into oxide. The resulting mold depression has a pyramid shape with predictable and precise geometry and small tip radius. Additionally, this anisotropic etch is somewhat self-limiting and is less therefore less sensitive to precise timing of the etching step. The entire description for this alternate process flow is presented in Appendix B.

VI.2. The Process Flow

Process 1: **Deposit Field Nitride**

A field nitride layer is deposited onto one side an undoped double-side polished wafer. This nitride dielectric serves as electrical isolation for the electrode pads.

- RCA clean
- 50:1 HF dip
- DI rinse and dry
- LPCVD of nitride (1 μm)

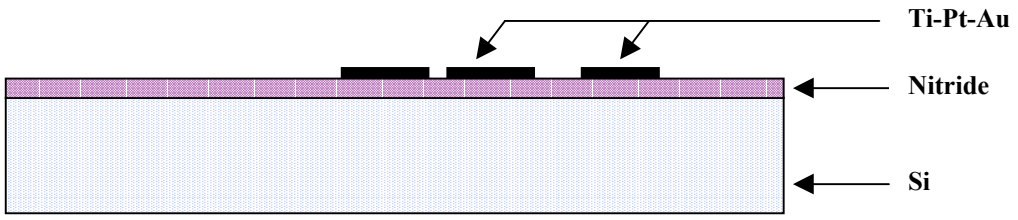


Process 2: **Mask #1 --- Metal Pads**

A Ti-Pt-Au metal tri-layer is deposited and patterned using a conventional liftoff process to create the actuation, self-test, and tunneling pads. The electrical contacts and conductive lines running to each pad are also fabricated. The tri-layer metal deposition is done to enhance adhesion of the gold layer to the wafer. The Pt layer is thermodynamically stable and prevents the migration of Ti through to the tunneling gold surface.

- RCA clean
- 50:1 HF dip
- DI rinse and dry
- HMDS, spin coat positive photoresist, and prebake
- Expose with Mask #1 ("Electrode Mask")
- Postbake and develop
- Deposit 50 Å Ti, and 50 Å Pt, and 400 Å Au

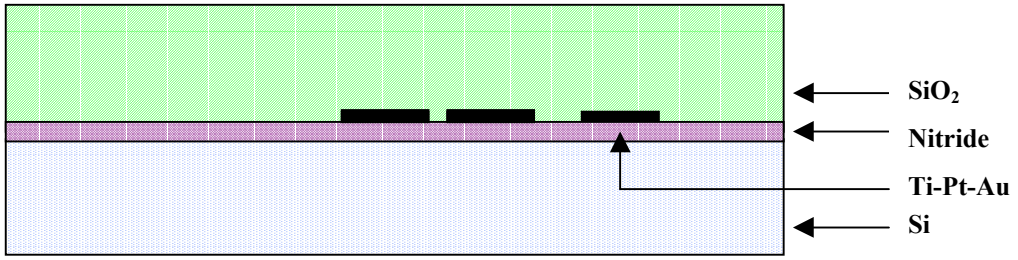
- Metal and photoresist liftoff by hot (40 C) ultrasonic acetone bath
- Nanostrip to clean off photoresist residue
- DI rinse and dry



Process 3: **Oxidation**

A sacrificial field oxide layer is deposited onto the wafer using low temperature plasma enhanced CVD. Low temperatures are desirable to maintain stability of the electrode metal.

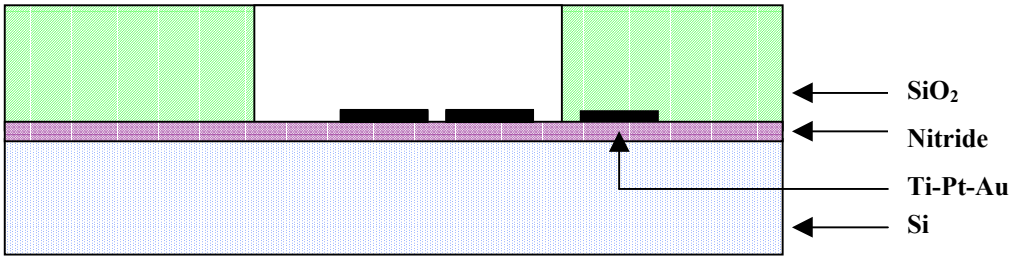
- Deposit 3.0 μm oxide by PECVD



Process 4: **Mask #2 --- Oxide Cavity**

A cavity that is nearly as large as the cantilever is etched into the oxide layer. This cavity is used to speed up the cantilever release process during the wet oxide etch of the sacrificial oxide.

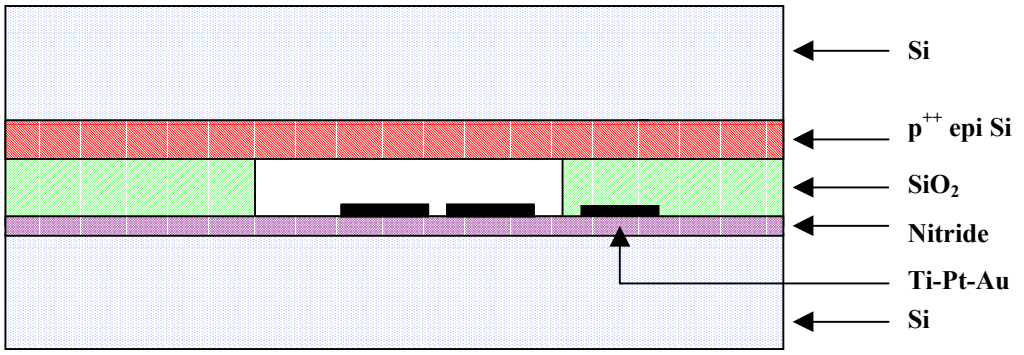
- HMDS, spin coat positive photoresist, and prebake
- IR alignment of Mask #2 ("Cavity Mask") and expose
- Postbake and develop
- 6:1 BOE etch oxide cavity
- DI rinse and dry



Process 5: **Planarize and Bond**

The oxide layer is CMP planarized and polished to a mirror-smooth finish. A second, undoped doped silicon wafer with a 2.0 μm layer of highly doped (> 10^{20} cm^{-3} boron) epitaxial silicon on one surface is obtained from a commercial vendor. The two wafers are then fused together along an oxide-epi bond line. No fine alignment procedures are required during the bonding process because the epi wafer has no patterned features.

- Polish (CMP) the oxide to 1.5 μm over Au
- Piranha clean
- Clamp wafers together and hold at 900 C, 1-4 barr, oxygen ambient
- Anneal at 800 C

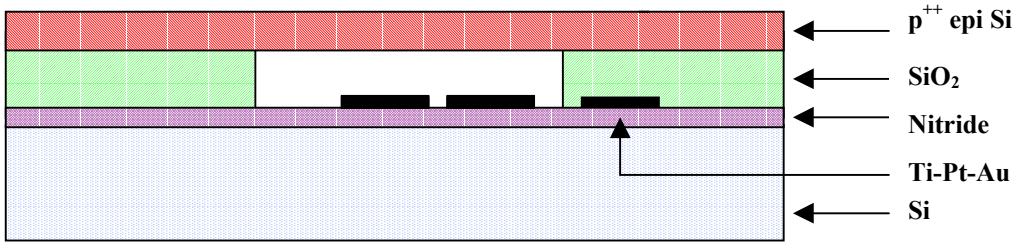


Process 6: **Thin Down to Etch-stop**

Gross wafer thinning of the upper bulk silicon layer is carried out by grinding. Precision thinning is achieved by an etch in EDP (ethylene diamine pyrocatechol) down to the p^{++} epi-Si etch stop. The etch rate of silicon with boron concentrations above 10^{20} cm^{-3} in EDP etchant is 350 times slower than for intrinsic silicon, so nearly all of the original two micron thickness of the epi-Si layer should be preserved after the thinning process is complete. At this point, the modified bonded SOI wafer is finished and ready for further processing.

- Backgrind top of the bonded wafer to 5 μm

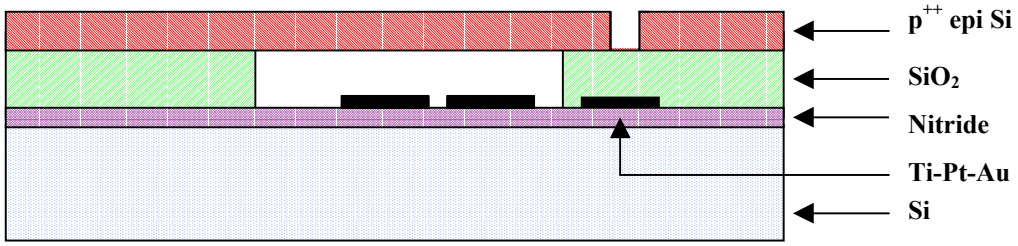
- EDP etch, 66 C
- Piranha and RCA clean



Process 7: **Mask #3 --- Etch Tip Hole Through epi-Si**

A mold for root of the tunneling tip is formed by first plasma etching a circular hole through the highly doped epitaxial layer.

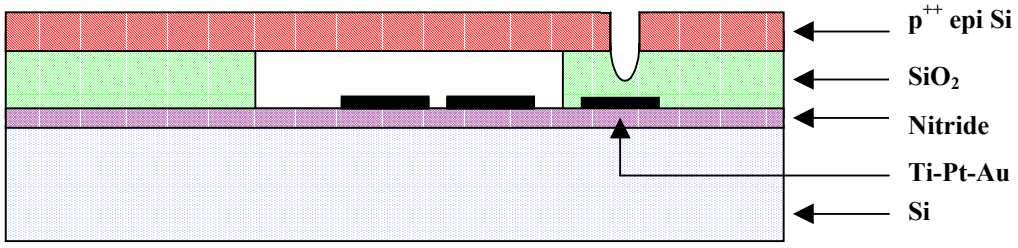
- RCA clean
- HMDS, spin coat positive photoresist, and prebake
- IR alignment of Mask #3 ("Tip Mold Mask") and expose
- Postbake and develop
- RIE etch hole through highly doped epitaxial Si
- Pirahna strip photoresist
- DI rinse and dry



Process 8: **Etch Tip Into Oxide**

The hole etched through the epi-Si layer in the previous process step is used as an etch mask for a timed partial etch of the underlying oxide layer. A spherical depression is isotropically etched into the oxide layer under high agitation. At this point the mold for the tunneling tip is complete.

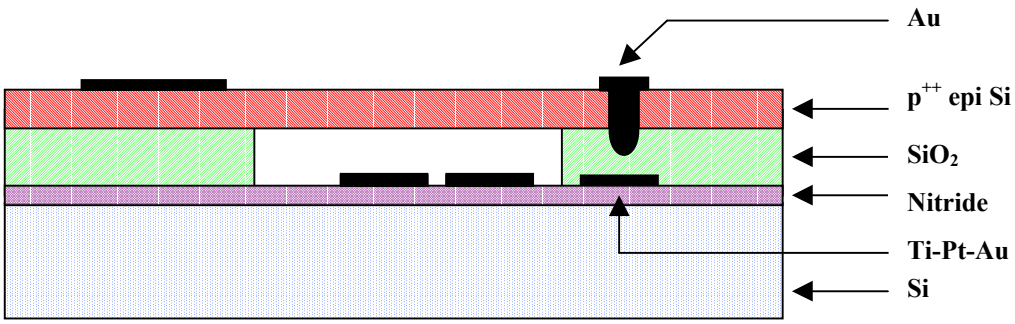
- Timed BOE
- Piranha strip photoresist
- DI rinse and dry



Process 9: **Mask #4 --- Metallize Tip and Contact**

A conducting tunneling tip is formed by deposition of gold into the depression previously etched into the oxide layer. The gold is patterned with a liftoff process. This process also forms the electrical contact pad to the highly doped epitaxial layer.

- HMDS, spin coat positive photoresist, and prebake
- Expose with Mask #4 ("Tip Metal Mask")
- Postbake and develop
- Deposit 10,000 Å Au
- Metal and photoresist liftoff by hot (40 C) ultrasonic acetone bath
- Nanostrip to clean off photoresist residue
- DI rinse and dry

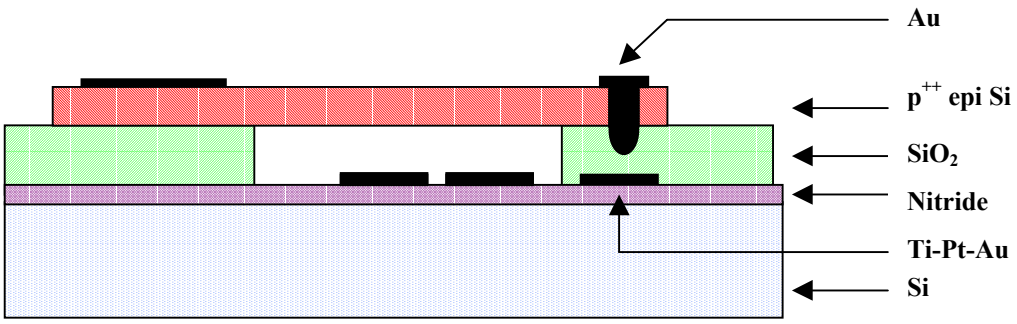


Process 10: **Mask #5 --- Define Cantilever**

The lateral dimensions of the epi-Si cantilever are defined using an RIE etch step. The tunneling tip and top contact are protected under photoresist during the etching process. The wafer is bonded to wafer mounting tape and the die is cut using a die saw. Finally, the photoresist is stripped.

- HMDS, spin coat positive photoresist, and prebake
- Expose to Mask #5 ("Cantilever Mask") and expose
- Postbake and develop
- RIE etch p^{++} epi-Si layer

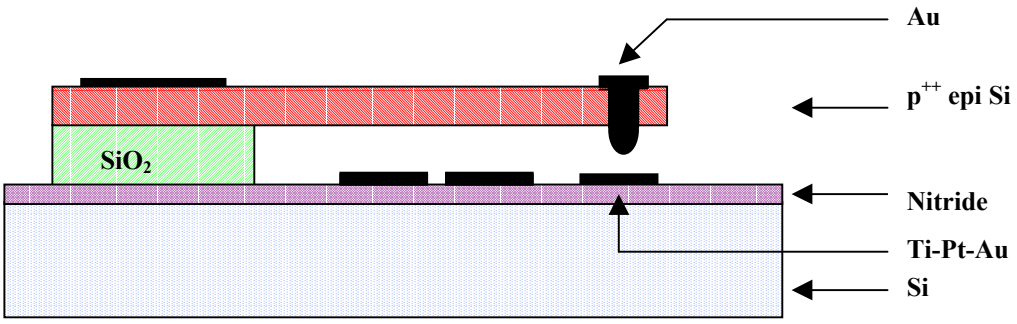
- Bond wafer to mounting tape
- Cut die with die saw
- Piranha strip photoresist
- DI rinse and dry



Process 11: Oxide Etch and Release

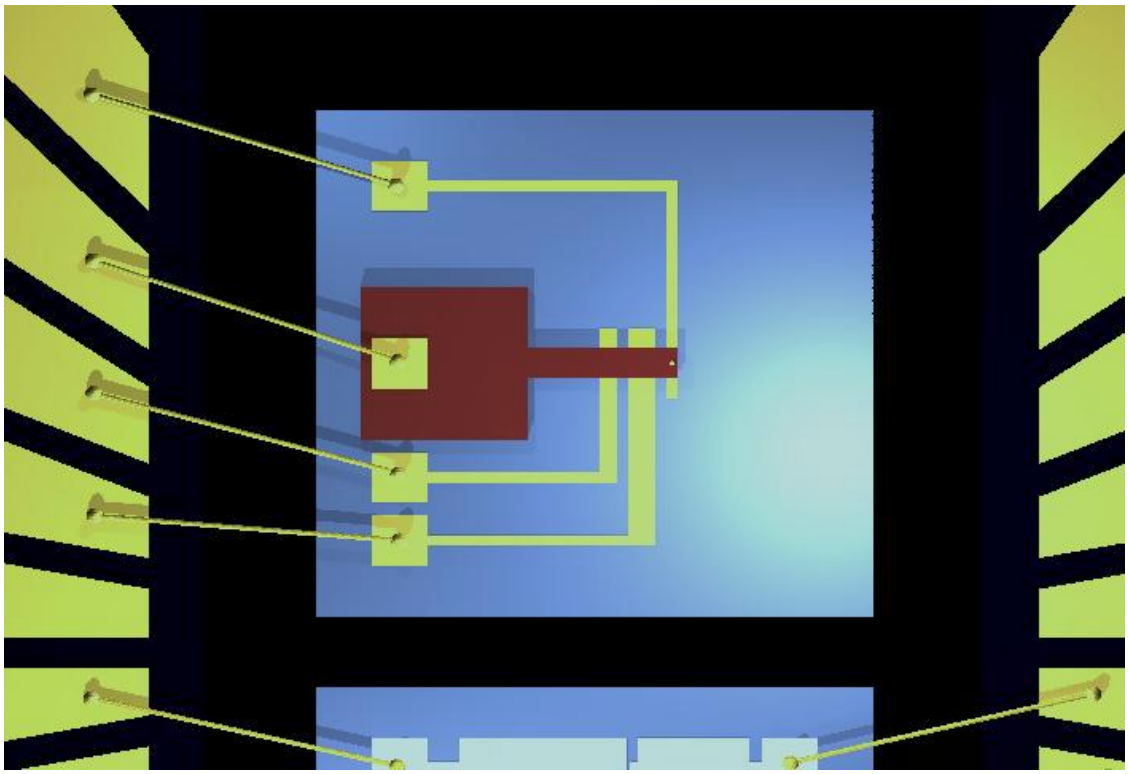
The cantilever is released during a timed wet etch of the sacrificial oxide layer. The time required for the wet etch is significantly reduced by the presence of the cavity previously etched into the oxide just below the cantilever. This short etch time also prevents significant undercutting of the oxide footing of the cantilever. It should be noted that no etch release holes are required to free the structure. At this time, the gold tunneling tip and counter electrode become exposed. Due to the long length of the cantilever relative to the gap, super critical CO_2 is employed to prevent adhesion of the cantilever to the floor of the gap.

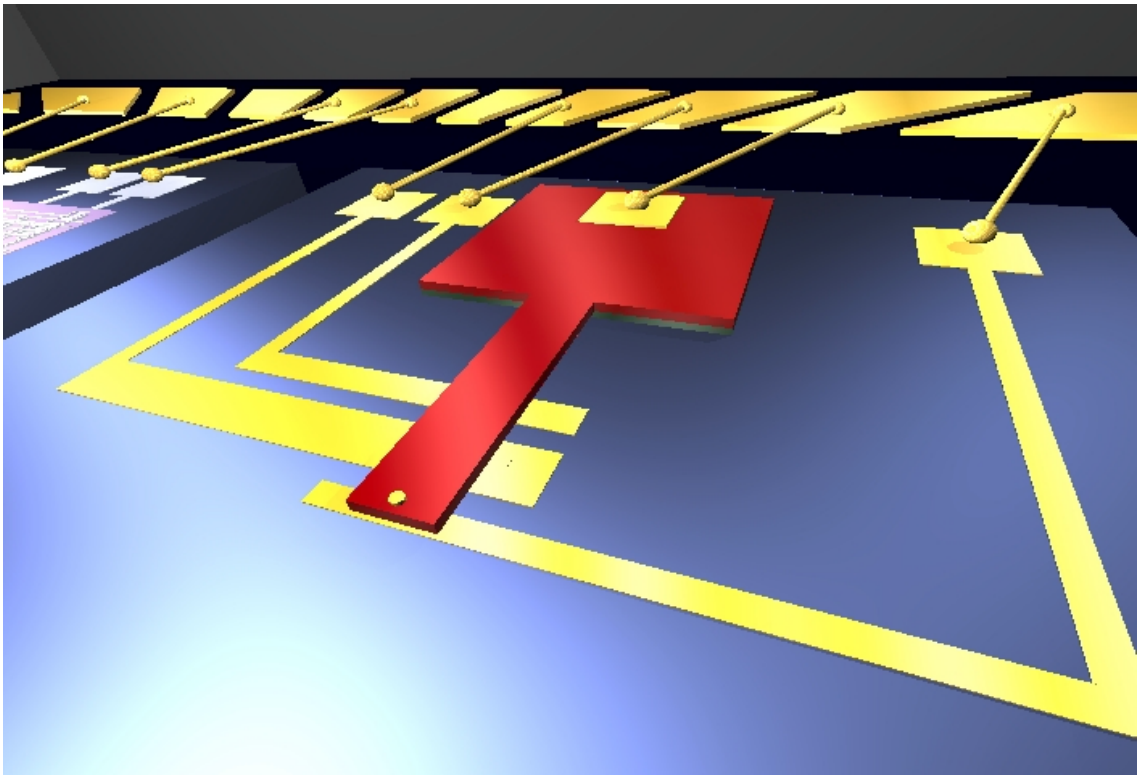
- Timed 6:1 BOE etch of the oxide layer
- Super critical CO_2 release

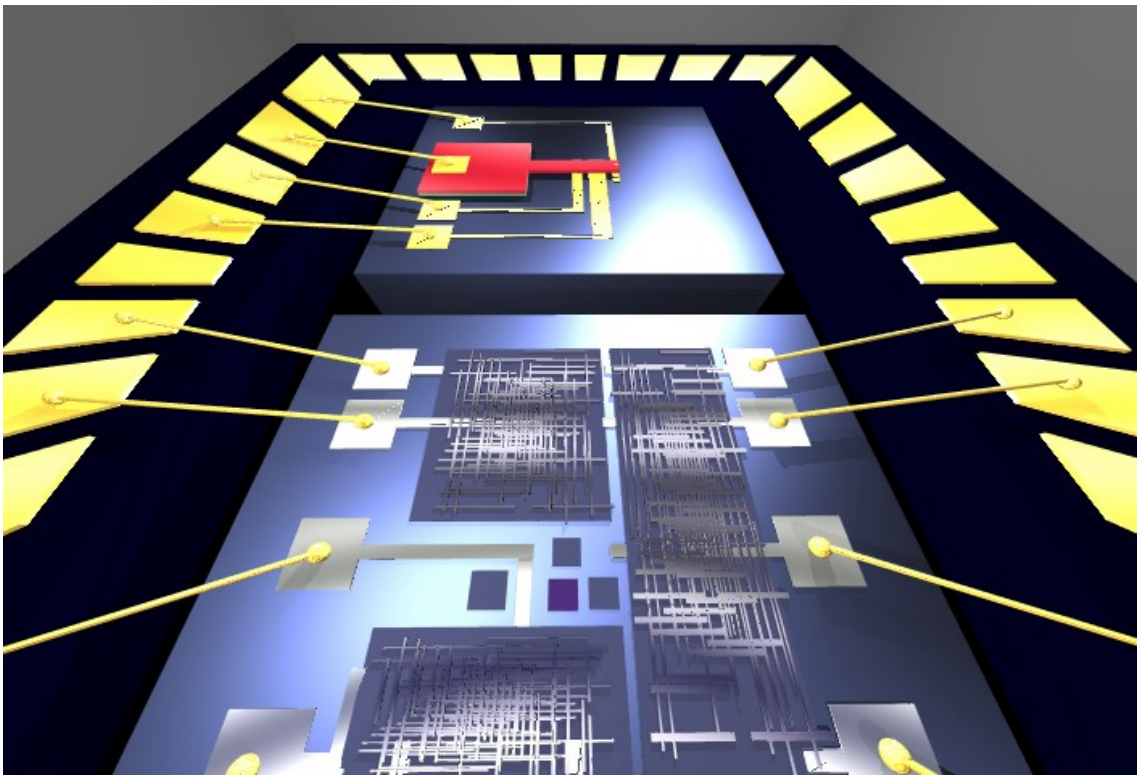


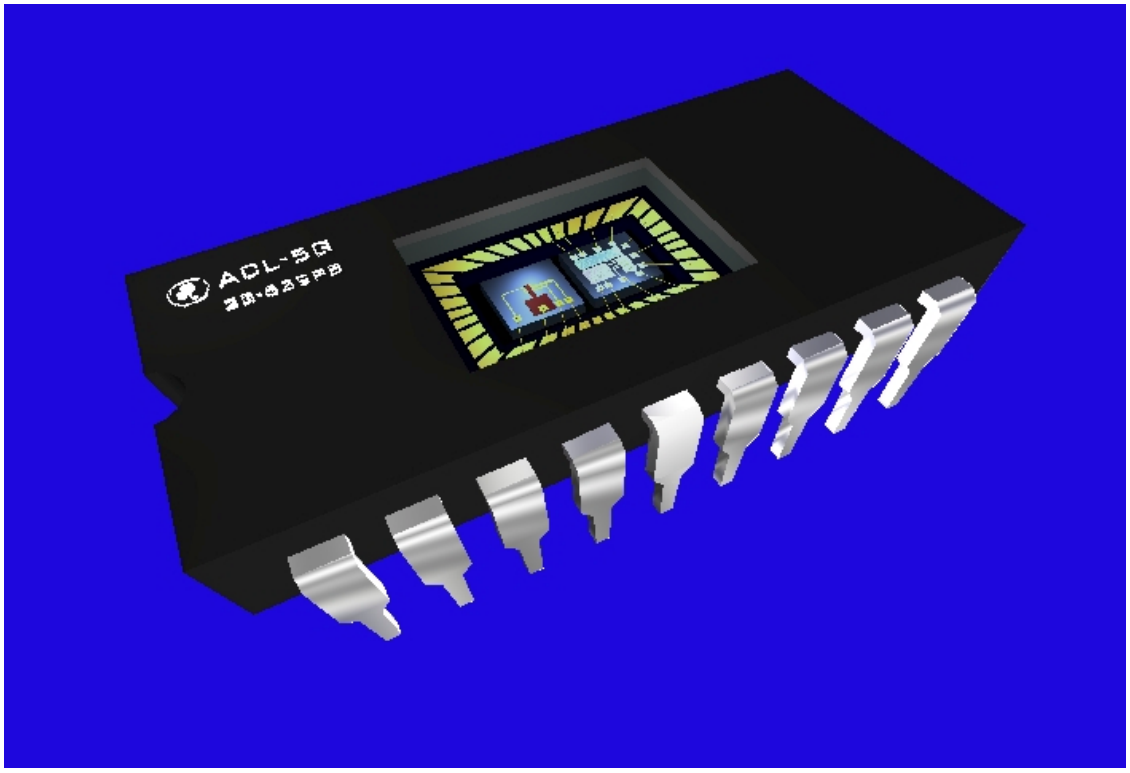
Process 12: Packaging

Finally, the cantilever chip and the electronics chip are mounted side-by-side in a two chip package, ball bonded, and sealed under nitrogen. Due the small size of both chips, through-hole dip-chip and subminiature surface mount configurations can be used.








ACL-5G

VII. Conclusion

The goal of this project was to design an accelerometer based on the operating principles of Scanning Probe Microscopy (SPM). The primary challenge was to fabricate the proof mass and the tip from the same wafer and still meet the following performance targets:

Full range	5g up to 100 Hz
Response time	10ms
Linearity	<1%
Cross-axis sensitivity	<1%
Self-test output	2.5g
Shock survivability	100g shock with 1ms risetime

As shown throughout this report, the model for the proposed tunneling accelerometer meets or exceeds every one of these specifications. In some cases, the performance benchmark exceeds the specification target by an order of magnitude or more:

Full range	5g up to 10,000 Hz
Response time	10 μs
Linearity	<<1%
Cross-axis sensitivity	<<1%
Self-test output	2.5g
Shock survivability	100g shock with <<1ms risetime

The success of the accelerometer model is due in no small part to the overall simplicity of the design. Whenever possible, we have made every effort to reject complexity in favor of simplicity. For example, many accelerometer designs in the literature use membranes with large suspended proof masses, doubly-supported beams, intracate tether structures, or complicated comb drives. However, our tunneling accelerometer is based on a canonical cantilever geometry. The single decision to use a simple cantilever structure has had profound and beneficial effect throughout our design process. Fabrication, control, sensitivity analysis, and dynamic modeling aspects of the design project have been simplified. At the same time, overall performance of the resulting device is significantly improved. In a similar fashion, the decision to use electron-tunneling sensing has also been beneficial. Unlike other accelerometer designs that use capacitive or piezoresistive sensing means which require large range of motion, our tunneling design uses only Angstrom-sized displacements. Because the displacements are so small, bulky and complex proof masses are not required. The smaller mass requirement gives the device higher bandwidth and improved cross-axis sensitivity.

This cantilever is easy to fabricate with traditional micromachining techniques. Very few mask steps are required. The thickness of the silicon and oxide layers that define the cantilever structure can be easily controlled to several tens of nanometers or less. This is fortunate because many of the performance characteristics of the accelerometer vary with the thickness of the cantilever to the third power and/or vary with the cantilever-electrode gap to the second power.

Also, because the cantilever is only supported at one end and is composed from one single-crystal material, it does not suffer from residual stresses caused by packaging or thermal expansion mismatch. The tunneling tip is simple to fabricate and is located at the intersection of the major axes of the cantilever. Therefore, torsional and off-axis vibrations are significantly reduced. Also, etch release holes are not used in the process flow. This makes modeling of dynamic effects like squeeze film damping a straight forward procedure.

At times, making things simple takes a lot of hard work. For example, much effort was devoted to proving that a the p^{++} epi-Si cantilever could be used as the conducting path for the tunneling tip as well as for the displacement path for the capacitive actuators. By using the cantilever as a single conductor, we save a mask step (which would have involve metallization or selective implantation of the cantilever layer) and avoid the complexities of trying to model a cantilever that is composed of two or more materials. And as a extra bonus, the highly doped epi-Si cantilever layer also serves as a precision etch stop when the modified SOI wafer is thinned down with anisotropic etchant.

These are just a few of the benefits of simple design.

APPENDIX A

This appendix contains the detailed tolerance analyses for the proof mass beam structure. Section A.1 presents a comparison of analytical and finite-element beam modal frequencies. Section A.2 presents calculations for the beam cross-axis sensitivities. Section A.3 presents 5% variational effects in the beam modal frequencies. Section A.4 presents results of more detailed beam geometric tolerance effects.

A.1. Behavior of a Cantilevered Beam

This section contains analytical expressions for the behavior of a beam, and compares the simple final geometry model analytical results to FEM results. Since the beam is quite wide, there is the possibility that beam theory may no longer be exact. As a result, the FEM analyses are used to determine how far from reality the analytical beam theory expressions are.

The expression for the endpoint deflection of a cantilevered beam in the z-direction due to an acceleration field, a, in the same direction is (Cra81):

$$\delta_{tip} = \frac{3L^4 \rho a}{2Et^2}$$

where L is the beam length, t is the beam thickness, ρ is the material density, and E is the material Young's modulus. Note that for ideal beam theory, the thickness, w, of the beam has no influence on the deflection. The "k/m ratio" of the beam in the z-direction is therefore:

$$\left(\frac{k}{m}\right)_z = \frac{a}{\delta tip} = \frac{2Et^2}{3L^4 \rho}$$

The "k/m ratio" of the beam in the y-direction is therefore (note that w and t reverse their meaning) (also note that this is the deflection in the y-direction due to an acceleration in that direction):

$$\left(\frac{k}{m}\right)_y = \frac{a}{\delta tip} = \frac{2Ew^2}{3L^4 \rho}$$

The modal frequencies of the beam in the z-direction and y-direction are:

$$f_z = \frac{1}{2\pi}\sqrt{\frac{\beta Et^2}{\rho L^4}}$$

$$f_y = \frac{1}{2\pi}\sqrt{\frac{\beta Ew^2}{\rho L^4}}$$

where β is a constant depending on the mode number (β=1.0302 for mode1, β=40.460 for mode2, β=317.219 for mode3, etc). The "k/m ratio" and modal frequency in the longitudinal x-direction (in response to an acceleration in the x-direction) are:

$$\left(\frac{k}{m}\right)_x = \left(\frac{E}{\rho}\right)\left(\frac{\pi^2}{4L^2}\right)$$

$$f_x = \left(\frac{1}{4L}\right)\sqrt{\frac{E}{\rho}}$$

With these analytical formulas, the behavior of the baseline design for the proof mass beam (described in Section 2) is analyzed. An ANSYS finite-element model iss also built and the modal frequencies are recorded. The boundary conditions in the finite-element model are matched to the ideal fixed end beam case. A comparison of these two analyses is shown in Table A1.

TABLE A1: Analytical Predicted Mode Frequencies vs. FEM Mode Frequencies

Mode Description	Analytical Prediction	ANSYS Model	Correlation
1st z-bending mode	40.281 kHz	40.748 kHz	GOOD
2nd z-bending mode	252.434 kHz	258.53 kHz	GOOD
1st torsion mode	-	381.81 kHz	-
3rd z-bending mode	706.830 kHz	765.40 kHz	GOOD
1st y-bending mode	1.208 MHz	> 1 MHz	GOOD
x-longitudinal mode	8.416 MHz	>> 1 MHz	GOOD

Bsed on the above results, it is clear that the analytical expressions for modal frequencies correspond very well with the exact solutions. Table A1 does not include an analytical prediction for the 1st torsion mode because a closed-form expression could not be verified. Using the ANSYS data, it is important to recognize the proximity of the 2nd z-axis bending mode and the 1st torsion mode to the 1st z-axis bending mode frequency. The 2nd z-axis modal frequency is 6.34 times the 1st z-axis modal frequency and the 1st torsion modal frequency is 9.37 times the 1st z-axis modal frequency. This separation is enough to justify the assumption in the full-system simulation that only the first mode frequency needs to be considered (see Section ?? - Osamah).

A.2. Cross-Axis Sensitivity

Based on the results in Section A.1, it is quite clear that the 1st y-axis bending mode and the 1st x-axis longitudinal mode occur at frequencies much higher than the 1st z-axis bending mode. The baseline directional "k/m ratios" (for the 1st modes in the x-, y-, and z-directions) are as follows:

$$\left(\frac{k}{m}\right)_z = 4.145x10^{10}\,s^{-2}$$

$$\left(\frac{k}{m}\right)_y = 3.731x10^{13}\,s^{-2}$$

$$\left(\frac{k}{m}\right)_x = 2.796x10^{15}\,s^{-2}$$

The "k/m ratio" for the y-axis behavior is 900.12 times that of the z-axis behavior and the "k/m ratio" for the x-axis behavior is 67455 times that of the z-axis behavior. As a result, these directional sensitivities are much less than the required 1% requirement of the z-axis sensitivity. Since the tunneling tip extends below the underside surface of the beam, calculations must be performed to verify that accelerations in the x and y-directions do not result in z-axis motion that is greater than 1% of the motion due to an equivalent acceleration in the z-direction. The mass of the tunneling tip is:

$$m_{tip} = \rho_{Si} L_{tip} A_{tip}$$

where L_{tip} is the tunneling tip length and A_{tip} is the tunneling tip area. For tip dimensions 2um x 2um x 1.3um, $A_{tip}= 4x10^{-12}$ m^2. An acceleration of 5g in the x-direction will act on the tip mass and produce a moment around the y-axis acting on the end of the beam:

$$M_{xaccel} = m_{tip} a \left(\frac{L_{tip}}{2} \right)$$

This moment acts to bend the beam upwards by an amount z_{xaccel}, according to the following relation:

$$z_{xaccel} = \frac{M_{xaccel} L^2}{2EI}$$

Therefore, the corresponding cross-axis sensitivity is:

$$\left(\frac{z_{xaccel}}{a} \right) = \frac{m_{tip} L_{tip} L^2}{4EI}$$

Substituting in the final geometry values of these variables, the following cross-axis sensitivity is obtained:

$$\left(\frac{z_{xaccel}}{a} \right) = 3.728x10^{-17} s^2$$

Recalling that the z-axis sensitivity for this final beam design is $1.16e^{-11}$ s^2, clearly the cross axis sensitivity is significantly less than 1% of the z-axis sensitivity. This previously calculated cross-axis sensitivity is for an x-direction acceleration of 5g. The other cross-axis sensitivity to be considered is that of a y-axis acceleration of 5g. It can argued, however, that this sensitivity will only be smaller than the above calculated one since the torsional mode frequency is larger than the 1st z-axis bending mode frequency. Therefore, in conclusion, this beam geometry possesses cross-axis sensitivities significantly less than 1% of the z-axis sensitivity.

A.3 Tolerance Analysis

Having shown in Section A.1 that the analytical expressions for "k/m ratio" and frequency are indeed valid, a 5% tolerance analysis is performed on the beam dimensions and material properties. The goal is to determine whether or not a 5% error in any of the dimensions or properties will have enough impact to bring the 2nd z-axis bending mode and 1st torsional mode down too close to the 1st z-axis bending mode frequency. These results are shown in Table A2, A3, A4, and A5. Analytical expressions are used to obtain the data in Tables A2, A3, and A4, whereas an optimization study is used to produce the data in Table A5. Each table simply states the % change in either "k/m ratio" or frequency for both +5% and -5% variation in the designated dimension or property. For example, in Table 3, when the beam length, L, is changed by -5%, the "k/m ratio" increases by a factor of 1.227 and the z-axis 1st modal frequency increases by a factor of 1.107.

TABLE A2: Z-axis 1st Modal Behavior

Parameter	k/m ratio + 5% parameter	k/m ratio - 5% parameter	frequency + 5% parameter	frequency -5% parameter
L	0.822	1.227	0.907	1.107
w	1	1	1	1
t	1.103	0.903	1.050	0.950
ρ	0.952	1.053	0.976	1.026
E	1.050	0.950	1.025	0.975

TABLE A3: Y-axis 1st Modal Behavior

Parameter	k/m ratio + 5% parameter	k/m ratio - 5% parameter	frequency + 5% parameter	frequency -5% parameter
L	0.822	1.227	0.907	1.107
w	1.103	0.903	1.050	0.950
t	1	1	1	1
ρ	0.952	1.053	0.976	1.026
E	1.050	0.950	1.025	0.975

TABLE A4: X-axis Longitudinal Mode Behavior

Parameter	k/m ratio + 5% parameter	k/m ratio - 5% parameter	frequency + 5% parameter	frequency -5% parameter
L	0.907	1.108	0.952	1.053
w	1	1	1	1
t	1	1	1	1
ρ	0.952	1.053	0.976	1.026
E	1.050	0.950	1.025	0.975

TABLE A5: Torsional Mode Behavior

Parameter	frequency + 5% parameter	frequency -5% parameter
L	0.947	1.051
w	0.980	1.047
t	1.047	0.947
ρ	0.976	1.026
E	1.025	0.975

Based on the equation for the z-axis bending modal frequencies, the ratio of the 2nd z-axis bending mode frequency to the 1st z-axis bending mode frequency will always be a constant equal to 6.267. The ratio of the 1st torsional mode frequency to the 1st z-axis bending mode frequency will vary, however, with certain dimensional variations. This ratio is unaffected by variations in ρ, E, and t, but is affected by variations in w and L. In fact, the worst case variation is a +5% variation in L coupled with a -5% variation in w. For this case, the 1st z-axis bending mode frequency increases from the baseline 40.748 kHz to 45.108 kHz and the 1st torsional mode increases from the baseline 381.81 kHz to 393.26 kHz. Therefore, the ratio decreases from 9.37 in the baseline case to 8.72 in the worst case. In conclusion, even with the worst case combination of parameter variation, the 2nd z-axis bending mode has a frequency at least 6.267 times greater than the 1st z-axis bending mode and the 1st torsional mode is always at least 8.72 times greater in frequency than the 1st z-axis bending mode.

A.4 Detailed geometric considerations

The finite-element model for the beam is now used to evaluate the effect of more detailed variations on the dynamic behavior of the beam structure. The critical mode frequencies to monitor throughout these studies are the 1st, ,2nd, and 3rd z-axis bending modes and the 1st torsional mode since these are the lowest four modes of the beam. The analyses in section A.1 and A.2 are based on a cantilever beam with a clearly known fixed boundary condition at the base end. In reality, this boundary condition may not be the case. Also, it could be that the beam dimensions themselves vary throughout the beam. The following analyses are carried out in order to evaluate these issues:

1) The size of the base structure to which the beam is attached (the thickness is fixed by the required gap, but the length and width are free variables to choose). If the base is too small, then the stiffness could be quite low, resulting in lower mode frequencies of the beam. In this analysis, bases of Si material are evaluated with dimensions 120um x 120um x 1.5um and 240um x 240um x 1.5um.
2) The material of the base between the Si beam and the substrate. This material is, in fact, SiO2. Again, since SiO2 has a lower Young's modulus (E=60GPa) than Si, the stiffness of the base could be reduced. In this analysis, bases of SiO2 are evaluated with dimensions 120um x 120um x 1.5um and 240um x 240um x 1.5um.
3) Thickness variation across the width of the beam. In this analysis, the thickness is varied from 0.95t to 1.05t across the width the beam.
4) Thickness variation along the length of the beam. In this analysis, the thickness is varied from 0.95t to 1.05t along beam toward the end in one case and from 1.05t to 0.95t along beam toward the end in another case.
5) The effect of tip mass on the mode frequencies. In this analysis, a tip with dimensions 5um x 5um x 1.3um is included on the underside of the beam at both the centerline of the width and at 10um off of the centerline of the width.

The results of these analyses are found in Table A6.

TABLE A6: Boundary Condition/Dimensional Asymmetry Issues

Analysis	Comments	z -axis 1st mode	z-axis 2nd mode	torsional 1st mode	z-axis 3rd mode
BASELINE	Fixed BC, nominal dimensions	40.748 kHz	258.53 kHz	381.81 kHz	765.40 kHz
1	Base: 60x60x1.5um	40.432 kHz	257.31 kHz	385.48 kHz	753.46 kHz
1	Base: 120x120x1.5um	40.452 kHz	257.63 kHz	379.71 kHz	775.29 kHz
2	Base: 60x60x1.5um	40.367 kHz	256.90 kHz	385.13 kHz	752.16 kHz
2	Base: 120x120x1.5um	40.416 kHz	257.40 kHz	379.56 kHz	774.47 kHz
3	0.95t-1.05t across beam	40.800 kHz	257.41 kHz	363.22 kHz	743.71 kHz
4	0.95t-1.05t along beam	38.297 kHz	253.38 kHz	347.54 kHz	715.35 kHz
4	1.05t-0.95t along beam	43.260 kHz	263.80 kHz	371.57 kHz	733.42 kHz
5	Centered tip mass	40.544 kHz	256.98 kHz	379.90 kHz	750.09 kHz
5	10um off-centered tip mass	40.561 kHz	257.34 kHz	382.96 kHz	746.50 kHz

Analyses 1 and 2 show that the presence of the SiO2 base material and the planar dimensions of the base itself have little, if any, effect on the beam modal frequencies. Therefore, for simplicity in modeling in all further analyses, the beam is forced to have fixed boundary conditions at the non-cantilevered end. This reduces finite-element computation time. Analysis 3 shows that variation in beam thickness across the width of the beam has little effect on the modal behavior of the beam. Analysis 4 shows that variation in beam thickness along the length of the beam does have some moderate effect (up to 5%) on the modal behavior of the beam. These are worst case fabrication scenarios and therefore, in reality would not be expected to happen. Analysis 5 shows that the presence of the small tunneling tip, even when offset from the centerline of the width by 10um, has little significant effect on the beam modal behavior. The first conclusion of these analyses is that a simplified model of the cantilever beam (not taking into account any base geometry or material) is sufficient to accurately predict the beam modal behavior. The second conclusion is that the beam modal behavior is unaffected by the presence of the tunneling tip and the beam is quite robust to variations in it thickness across the width and along the length.

APPENDIX B

This appendix contains an alternate process flow for the fabrication of a tunneling accelerometer with a tunneling tip that has more precise geometry and smaller tip radius. This process flow is similar to the process flow presented in the main portion of this paper. However, the primary modification is that the mold for the tunneling tip is anisotropically etched in to an undoped epi-silicon layer instead of isotropically into oxide. The resulting mold depression has a pyramid shape with predictable and precise geometry and small tip radius. Additionally, this anisotropic etch is somewhat self-limiting and is less therefore less sensitive to precise timing of the etching step.

B.1. Overview

We have developed a four mask process flow for the fabrication of the microelectromechanical assembly for a tunneling accelerometer with a precision fabricated tunneling tip. A cantilever proof mass with an integrated tip is created within a single bonded wafer. The dimensions of the completed proof mass are $270 \times 60 \times 2$ μm with a nominal 1.5 μm gap between the actuation, self-test, and sense electrodes and the underside of a conductive p^{++} epi-Si cantilever. The tunneling tip protrudes from the underside of the cantilever into the gap between the cantilever and electrode pads.

The process flow described in this section satisfies four major challenges. The first challenge was set forth in the initial design problem statement. Specifically, the tunneling tip and the proof mass must are fabricated from the same wafer. The squeeze film damping requirements of accelerometer dictated the second challenge. Etch release holes in the beam drastically reduce the squeeze film damping effects. Therefore, the cantilever is released without the aide of etch release holes. Thirdly, a tip of precise and predictable geometry is fabricated onto the underside of the cantilever beam. The fourth challenge was self-imposed by the design team. Specifically, every effort has been made to design a simple process flow that uses the fewest number of masks or process steps, and does not require unusual or expensive fabrication technologies. The expected benefits of meeting this third challenge include less uncertainty in the fabrication process, improved tolerance benchmarks, reduced time to market, and lower cost per die, among others.

The starting materials consist of two wafers. Wafer "A" is an undoped double-side polished wafer. First, a nitride dielectric layer is deposited for electrical isolation. Second, gold metal (with two additional adhesion promoting metals) is deposited and patterned to form the electrodes and contact pads. Next, an oxide layer is deposited. Then a cavity is etched into the oxide layer just below the point where the cantilever is to be formed. This presence of this cavity significantly reduces the lateral dimension of oxide that must be removed during the final cantilever release process. Therefore, etch release holes are not required and the second challenge is satisfied. Wafer "B" is a commercially prepared undoped doped silicon wafer with a 2.0 μm layer of highly doped ($> 10^{20}$ cm^{-3} boron) epitaxial silicon on one surface. A undoped epitaxial silicon layer is grown on top of the highly doped epitaxial silicon layer of Wafer "B", planarized, and polished. The two wafers are bonded together along an oxide-epi bond line without the use of special alignment procedures. The bonded wafer is thinned down to the highly

doped epitaxial layer by grinding and etching. The tunneling tip is formed by first etching a hole through the highly doped epitaxial layer. Then a pyramid-shaped depression is anisotropically etched into the undoped epitaxial layer. The depression is used as mold for subsequent deposition of gold. This forms a pyramid-shaped gold tunneling tip of precise geometry and small tip radius. Then lateral dimensions of the cantilever are defined, metal contact pads are deposited, and the cantilever is released. Finally, the undoped epitaxial layer on the bottom side of the cantilever is removed, thereby revealing the tunneling tip.

B.2. The Process Flow

Process 1: Isolation and Electrodes

A field nitride layer is deposited onto one side of undoped double-side polished Wafer "A". This nitride dielectric serves as electrical isolation for the electrode pads. A Ti-Pt-Au metal tri-layer is deposited and patterned onto double-side polished Wafer "A" with a liftoff process to create the AC, DC, self-test, and tunneling pads. The electrical contacts and conductive lines running to each pad are also fabricated. The tri-layer metal deposition is done to enhance adhesion of the gold layer to the wafer. The Pt layer is thermodynamically stable and prevents the migration of Ti through to the tunneling gold surface.

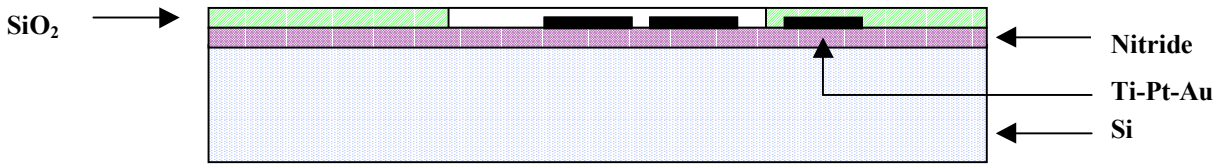
- RCA clean
- 50:1 HF dip
- DI rinse and dry
- LPCVD of nitride (1 μm)
- HMDS, spin coat positive photoresist, and prebake
- Expose with Mask #1 ("Electrode Mask")
- Postbake and develop
- Deposit 50 Å Ti, and 50 Å Pt, and 400 Å Au
- Metal and photoresist liftoff by hot (40 C) ultrasonic acetone bath
- Nanostrip to clean off photoresist residue
- DI rinse and dry

Process 2: Sacrificial Oxide Layer

A sacrificial field oxide layer is deposited conformally onto Wafer "A". A cavity that is nearly as large as the cantilever is etched into the oxide. This cavity is used to speed up the cantilever release process during the wet oxide etch of the sacrificial oxide. Finally, the oxide layer is thinned, planarized, and polished to a mirror-smooth finish.

- Deposit 1.5 μm oxide by PECVD
- HMDS, spin coat positive photoresist, and prebake
- Expose with Mask #2 ("Cavity Mask")
- Postbake and develop

- 6:1 BOE etch oxide cavity
- DI rinse and dry
- Polish (CMP) the oxide to 0.3 μm over Au
- Piranha clean



Process 3: Epitaxial Layer

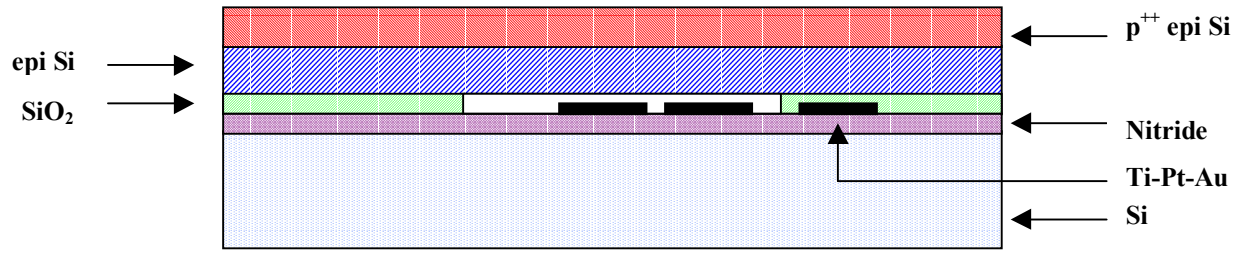
Undoped epitaxial silicon is grown on top of the 2.0 μm highly doped epitaxial layer of Wafer "B". This sacrificial layer will be used as a mold to precisely control the geometry of the tunneling tip in subsequent fabrication steps. Finally, the epitaxial layer is thinned and planarized and polished to a mirror-smooth finish.

- RCA clean
- 50:1 HF dip
- Grow 3.5 μm undoped epitaxial Si
- Polish (CMP) the undoped epitaxial layer to 1.2 μm
- Piranha and RCA clean

Process 4: Wafer Bonding and Thinning

Wafers "A" and "B" are bonded together along an oxide-epi bond line. No special alignment procedures are required because Wafer "B" has no patterned features. Gross wafer thinning is carried out by grinding. Precision thinning is achieved by an etch in EDP (ethylene diamine pyrocatechol) down to the highly doped epitaxial layer etch stop. The etch rate of silicon with boron concentrations above 10^{20} cm^{-3} in EDP etchant is 350 times slower than for intrinsic silicon.

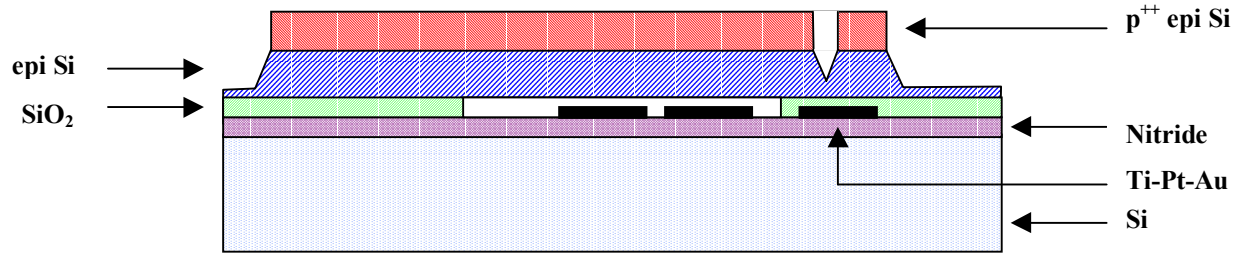
- Clamp wafers together and hold at 900 C, 1-4 barr, oxygen ambient
- Anneal at 800 C
- Backgrind top of the bonded wafer to 5 μm bulk Si
- EDP etch, 66 C
- Piranha and RCA clean



Process 5: Tip Mold and Cantilever Definition

A mold for the tunneling tip is formed by first plasma etching a square hole through the highly doped epitaxial layer. Then a pyramid-shaped depression is anisotropically etched with KOH into the undoped epitaxial layer. The <100>/<110> etch rate ratio for intrinsic silicon in KOH is 400 to 1, which provides a depression with a narrow tip and sloping side walls. There are no etch stops so careful timing is desirable. These two etch processes also help to define the lateral dimensions of the highly doped epitaxial cantilever and its supporting root.

- RCA clean
- HMDS, spin coat positive photoresist, and prebake
- Expose with Mask #3 ("Tip Mold Mask")
- Postbake and develop
- RIE etch hole through highly doped epitaxial Si, no etch stop
- KOH etch pyramid depression in undoped epitaxial Si, no etch stop
- Pirahna strip photoresist
- DI rinse and dry

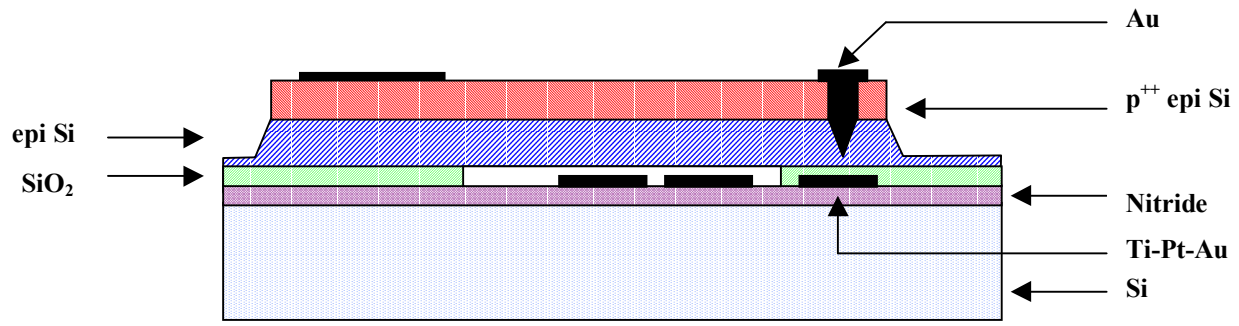


Process 6: Tunneling Tip

A conducting tip is formed by deposition of gold into the pyramid-shaped depression previously etched into the undoped epitaxial layer. The gold is patterned during a liftoff process. This process also forms the electrical contact pad to the highly doped epitaxial layer.

- HMDS, spin coat positive photoresist, and prebake

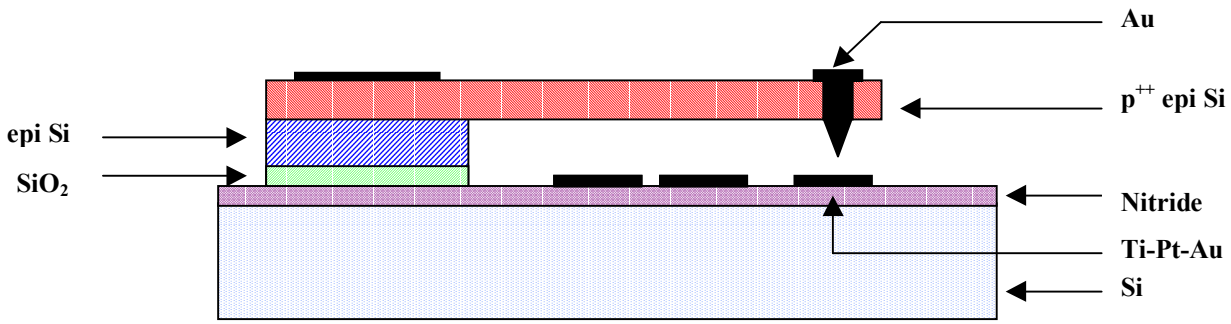
- Expose with Mask #4 ("Tip Metal Mask")
- Postbake and develop
- Deposit 10,000 Å Au
- Metal and photoresist liftoff by hot (40 C) ultrasonic acetone bath
- Nanostrip to clean off photoresist residue
- DI rinse and dry



Process 7: Cantilever and Tip Release

A timed EDP dip is used to clear away the fraction of a micron of field undoped epitaxial silicon that still remains on the wafer. The cantilever is released during a timed wet etch of the sacrificial oxide layer. The time required for the wet etch is significantly reduced by the presence of the cavity previously etched into the oxide just below the cantilever. Finally, the 1.2 μm undoped epitaxial silicon layer on the under side of the released cantilever is selectively removed by EDP etchant. Because the oxide layer has be previously removed, the EDP etchant has free access to the entire region of undoped epi-Si on the underside of the cantilever. This second EDP etch process exposes the gold tunneling tip, reduces the thickness of the cantilever to 2.0 μm, and removes bulk intrinsic silicon between the electrodes on the floor on the air gap. The 350:1 selectivity of EDP for undoped versus doped silicon protects the lateral dimensions of the cantilever during both EDP etch processes. The completed tunneling tip protrudes from the underside of the cantilever and has a small tip radius and pyramidal geometry.

- Timed EDP etch of the exposed undoped epitaxial layer, 66 C
- Timed 6:1 BOE etch of the oxide layer
- EDP etch of the undoped epitaxial layer and bulk intrinsic silicon, 66 C
- Super critical CO_2 release


Au
p++ epi Si
epi Si
SiO2
Nitride
Ti-Pt-Au
Si

References

(Ana) Analog Devices Web Site. http://www.analog.com.

(Ash95) Ashburn, Michael, *A High Precision Fully Integrated Accelerometer*, MIT S.M. Thesis, October 1995.

(Che93) Chen, C. Julian, *Introduction to Scanning Tunneling Microscopy*. Oxford UP, New York. (1993)

(Cra81) Craig, R., *Structural Dynamics: An Introduction to Computer Methods*. John Wiley & Sons, New York. (1981)

(Deb90) de Bruin, D,et al., "Second-order effects in Self-testable Accelerometers," IEEE Solid-State Sensor and Actuator Workshop, June 4-7, 1990, pp149-152.

(Hor89) Horowitz, P., Hill, W., *The Art of Electronics*. Cambridge UP. (1989)

(Kai) Kaiser, W., Jaklevic, R., *IBM J. Res. Develop*. **30**:411. (1986)

(Ken) Kenny, T.W., Kaiser, W.J., Rockstad, H.K., Reynolds, J.K., Podosek, J.A., Vote, E.C., "Wide-bandwidth electro-mechanical actuators for tunneling displacement transducers." Center for Space Microelectronics Technology, Jet Propulsion Laboratory, California Institute of Technology.

(HCP) From *Handbook of Chemistry and Physics*, 69th edition, CRC Press (1988).

(How93) Howland, R., Benatar, L., *A Practical Guide to Scanning Probe Microscopy*. Park Scientific Instruments. (1993)

(Wie94) Wiesendanger, R., *Scanning Probe Microscopy and Spectroscopy*. Cambridge UP, Cambridge. (1994)

(Yeh98) Yeh, C., Najafi, K., "CMOS Interface Circuitry for a Low-voltage Micromachined Tunneling Accelerometer." *Journal of Microelectromechanical Systems*, **7**.1:6-15. (1998)

Bibliography

1. R.P. van Kampen, *Bulk-Micromachined Capacitive Servo-Accelerometer*, Delft University Press, 1995.
2. Mitchell J. Novack, *Design and Fabrication of a Thin Film Micromachined Accelerometer*, MIT PhD Thesis, September 1992.
3. Philip W. Barth, et al, "A Monolithic Silicon Accelerometer With Integral Air Damping and Overange Protection," IEEE Solid-State Sensor and Actuator Workshop, June 1988, pp35-38.
4. Daniel Lapadatu, Michel De Cooman, Robert Puers, "A Double Sided Capacitive Miniaturised Accelerometer Based on Photovoltaic Etch Stop Technique," The 8th International Conference on Solid-State Sensors and Actuators, Stockholm, Sweden, June 25-29, 1995, pp546-549.
5. Yoshinori Matsumoto et al., "A Capacitive Accelerometer Using SDB-SOI Structure," ," The 8th International Conference on Solid-State Sensors and Actuators, Stockholm, Sweden, June 25-29, 1995, pp550-553.
6. Howard K. Rockstad et al., "A Miniature, High-Sensitivity, Electron Tunneling Accelerometer," ," The 8th International Conference on Solid-State Sensors and Actuators, Stockholm, Sweden, June 25-29, 1995, pp674-678.
7. Minhang Bao, et al., "A Micromechanical Structure Eliminating Lateral Effect of Silicon Accelerometer," Transducers 91 1991 International Conference on Solid-State Sensors and Actuators, June 24-28, 1991, pp101-103.
8. E. Peeters et al., "A Highly Symmetric Capacitive Micro-Accelerometer With Single Degree of Freedom Response," Transducers 91 1991 International Conference on Solid-State Sensors and Actuators, June 24-28, 1991, pp97-100.
9. Jun-Hwan Sim, et al., "Eight-beam piezoresistive accelerometer fabricated by using a selective porous-silicon etching method," Sensors and Actuators A 66 (1998), pp273-278.
10. Kijin Kwon and Sekwang Park, "A bulk-micromachined three-axis accelerometer using silicon direct bonding technology and polysilicon layer," Sensors and Actuators A 66 (1998), pp250-255.
11. Cheng-Hsien Liu, "Characterizaton of a High-Sensitivity Micromachined Tunneling Accelerometer with Micro-g Resolution," Journal of Microelectromechanical Systems, Vol 7, No 2, June 1998, pp235-243.
12. Charles Heng-Yuan Hsu, Silicon Microaccelerometer Fabrication Technologies, MIT PhD Thesis, September 1997.