Analog VLSI for Active Drag Reduction

Thesis by

Vidyabhusan Gupta

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1997

(Submitted January 24, 1997)

Acknowledgements

The process of obtaining a Ph.D. is often secondary to the knowledge and experience that one acquires along the way. One of the best things that can be said about my graduate study at Caltech are the wonderful interactions and friendships that I have developed. Many people have assisted me through the course of my life and, in particular, my graduate studies and I would like to acknowledge them.

First, I would like to thank my parents, Vijay and Uma Gupta, without whose life-long support I would not have been able to follow my muse of electrical engineering which led me finally to graduate school at Caltech.

I would like to also acknowledge the support given to me by my advisor, Professor Rodney Goodman. I have really enjoyed the academic freedom and support available in Rod's lab. I must also acknowledge the national Science Foundation and the Intel Foundation for their graduate fellowship support. Support from these organizations has helped further my graduate career.

Several other professors have strongly facilitated my learning experience. Professor Chih-Ming Ho and Professor Yu-Chong Tai have been instrumental in providing guidance and facilities to use for my research. Professor Carver Mead and his whole lab opened my eyes to the power of analog computation.

Fukang Jiang, Dr. Steve Tung, Tom Tsao, Chang Liu, and Charles Grosjean have been my fellow researchers on the quest to build the active skin. Without their invaluable help and knowledge, I would not have been able to finish.

My fellow lab members John Miller, Andy Moore, John Cortese, David Babcock, Jeff Dickson, and Vincent Koosh gave me a willing ear when I needed it. A special thanks to Dave who gave me the most cogent feedback on this thesis.

To the extended Caltech analog VLSI design community consisting of many friends like Kwabena (Buster) Boahen, Doug Kerns, Gert Cauwenberghs, and Tobi Delbrück, I can just simply thank them for all the help and guidance they have

offered me over the years. I have also been lucky to interact over the years with Art Zirger and Bob Freeman whose friendships are valued highly.

Although my wife Anirma is a recent ray of sunshine in my life, she had a deep impact on my work. Her patience, understanding, and guidance are acknowledged and deeply appreciated.

Finally, to Sagar, my dear brother, with whom I am unable to share the conclusion of work he witnessed, I say that I hope to live up to your ideals of hard work and success one day.

Abstract

In today's cost-conscious air transportation industry, fuel costs are a substantial economic concern. Drag reduction is an important way to increase fuel efficiency which reduces these costs. Even a 1% reduction in drag can translate into estimated savings of tens of millions of dollars in annual fuel costs.

Fluid mechanicists believe that microscopic vortex pairs impinging on the surface play an important role in turbulent transport that may cause large skin friction drag. The microscopic nature and unpredictable appearance of these structures has limited practical approaches to their control. With the advent of micromachining technology providing the ability to build mechanical structures with microscopic dimensions, the tools finally exist with which to detect and control the vortex structures. These sensors and actuators require control circuitry between them in order to build a complete system.

We propose an analog VLSI system that can process information along a surface in a moving fluid with the goal of controlling actuators to minimize the surface shear stress. We obtain the information from the surface by using microsensors which measure the surface shear stress. The actuators interact with the fluid by moving up and down in an attempt to diminish the impact of the drag-inducing structures in the fluid.

We have designed and fabricated and analog control system. We have tested the system in several different experiments to verify its effectiveness in providing a control signal that energizes an actuator. We also have studied the methodology for creating a completely integrated wafer-scale system.

Contents

List of Figures

List of Tables

Introduction

In this thesis we present an analog very large scale integration (VLSI) approach to the problem of turbulent drag reduction. The control of this drag is a highly sought after goal because large amounts of fuel energy must be expended to counteract it. It is estimated that about 30% - 50% of the total drag of an aircraft is produced by turbulent skin friction [21].

In 1988, the fuel bill for commercial aviation in the United States totaled approximately $10 billion. Reducing the drag by just 1%, would translate into an annual fuel savings of $50 million. This is certainly a substantial amount of money to be saved and is a strong motivating factor behind the research into understanding and controlling turbulent drag [21].

Researchers have speculated that interesting and viable techniques exist for the control of the phenomena of turbulent drag [23]. In a fluid mechanical sense, the problem has been identified for many years in one formulation or another. Only recently with better wind tunnel instrumentation and especially numerical computer-aided simulations of the phenomena, have researchers been able to begin understanding the mechanisms and possible solutions to turbulent drag.

One phenomena, known as counter-rotating longitudinal vortices, occurs in a turbulent flow over a surface and is thought to be responsible for a large portion of the drag. The microscopic size of these vortices, which decreases as the Reynolds number of the flow increases, has limited physical experimentation. Additionally, the inherent complexity of the non-linear Navier-Stokes equations has likewise limited analytical approaches.

With the advent of new microscopic mechanical devices, we can now begin to tackle the problem of building a drag-reducing system. We can fabricate structures that interact with the fluid at the necessary length scales.

Our approach is to develop a basic understanding of the fluid mechanical issues

and then to look to biological systems for guidance in selecting a methodology to use. We would like to utilize available technologies and algorithms to build the system.

System Development

The system we wish to build will react to a turbulent fluid moving over it. We need sensors to measure the surface shear stress and actuators to influence the flow. The key component between the sensors and actuators is the signal processing system which implements the appropriate control algorithm. Figure 0.1 is a very high level approach to the system where we depict how the different blocks interact with each other. This thesis will describe this signal processing system.

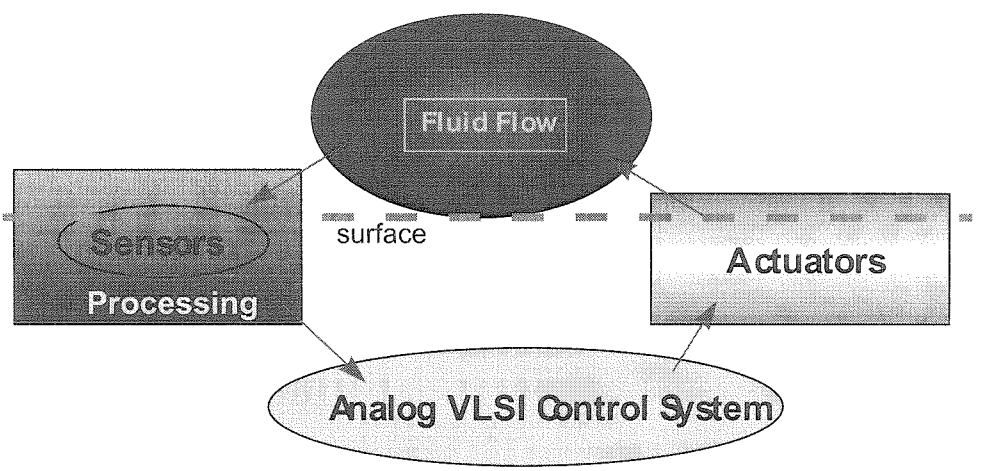

Figure 0.1: A high level view of the proposed system depicting the major components. The sensors allow observations of the fluid flow and the actuators affect the flow. The analog VLSI circuit between the sensors and actuators performs the signal processing to implement a controller.

Since we will use silicon micromachining technology to build the sensors and actuators and CMOS technology for the analog circuits, some thought is given in how to integrate the two. In the course of this research we have also investigated the issues involved in this important integration of CMOS electronics and silicon micromachining structures.

Chapter Summary

In Chapter 1 we summarize the basics of fluid mechanics. The purpose of this chapter is to give the reader a general overview of the issues involved in turbulent drag and the corresponding statistics.

In Chapter 2 we examine one of nature's more interesting systems that offers insight to the possibility of achieving drag reduction.

In Chapter 3 we look at drag reduction schemes that other people have employed and we touch on some neural network research that yields some clues of how the problem might be attacked.

In Chapter 4 we give a brief overview of the micro-electromechanical system (MEMS) technology employed in building a viable system. Especially important are the electrical properties of the μ-shear stress sensors and μ-actuators. Knowledge of these properties aids in the development of the necessary interface circuits as well as guides algorithm development.

In Chapter 5 we explain the information processing system that we have built and tested. We tie together the basis for the system as explained in the previous chapters and explain how the system performs turbulence detection. In particular, we give details for the operation of the individual circuits and the methodology for connecting the circuits together.

In Chapter 6 we describe the results obtained from testing the system in a variety of experiments. The experiments range from simple chip testing on the lab bench to capturing signals resulting from turbulent flow experiments in the wind tunnel.

In Chapter 7 we detail the experiments and procedures that we have used to build a completely integrated system. We present a number of methods used to build MEMs devices on the same substrate as CMOS circuits. A number of technical issues exist that need to be addressed in order to achieve the desired level of integration between the MEMs and CMOS circuitry.

Part I

Background

Chapter 1 Fluid Mechanics

Fluid flow over surfaces has been studied by fluid mechanicists for many years. The interaction between the moving fluid and the surface is studied in many ways. One method is to experimentally measure the interaction between the fluid flow and a surface in a wind tunnel. Another example is to study the problem computationally via computer simulations using approximate solutions to the basic laws of fluid motion.

When a fluid flows over a surface, the surface experiences a force due to the movement of the fluid. The force parallel to the direction of fluid movement is called drag and consists of two components. One component is called skin friction drag, the stream-wise component of the integral of all shearing forces on the surface. The other component is called pressure or form drag, the integral of all normal forces on the surface.

$$\text{total drag} = \text{skin friction drag} + \text{pressure drag} \tag{1.1}$$

1.1 Fluid Mechanics Notation

Throughout this thesis we will refer to various standard quantities in fluid mechanics. Table 1.1 and Table 1.1 presents these quantities. Figure 1.1 illustrates the coordinate system to help orient the reader.

1.2 Fluid Transport

When a fluid flows over a surface, two general regions can be identified. The outer region, or free-stream flow, is situated away from the surface well in the fluid. This

Variable	Description	Units
x	streamwise coordinates	m
y	wall normal coordinates	m
z	spanwise coordinates	m
t	time	s
U_c	centerline wind tunnel airspeed	$\frac{\text{m}}{\text{s}}$

Table 1.1: A listing of the basic variables of measurement in fluid mechanics.

Derived Variable	Description	Derivation	Units
u	streamwise velocity	$\frac{\partial x}{\partial t}$	$\frac{\text{m}}{\text{s}}$
v	wall normal velocity	$\frac{\partial y}{\partial t}$	$\frac{\text{m}}{\text{s}}$
w	spanwise velocity	$\frac{\partial w}{\partial t}$	$\frac{\text{m}}{\text{s}}$
τ_w	wall shear stress	$\mu \frac{\partial u}{\partial y}\|_w$	Pa
U_τ	wall shear or friction velocity	$\sqrt{\frac{\tau_w}{\rho}}$	$\frac{\text{m}}{\text{s}}$
u^+	non-dimensional streamwise velocity	$\frac{u}{U_\tau}$	
λ^+	viscous length scale	$\frac{\nu}{U_\tau}$	m
y^+	wall normal unit scale	$\frac{U_\tau y}{\nu}$	
x^+	wall stream-wise unit scale	$\frac{U_\tau x}{\nu}$	
z^+	wall span-wise unit scale	$\frac{U_\tau z}{\nu}$	
t^*	turbulent time scale	$\frac{y}{U_\tau}$	

Table 1.2: A tabulation of commonly used derived variables in fluid mechanics and their derivations.

region is considered inviscid[1]. The inner region, close to the surface, is known as the boundary layer. It is this boundary layer that is important for the problem of drag reduction. Unlike the free stream region, the boundary layer is considered a viscous fluid region. The appearance of the boundary layer is related to the fact that the tangential velocity of the flow at the surface is zero. This zero velocity comes about because the equations of Newtonian fluid movement impose a no-slip condition at the surface. The velocity increases towards its free stream value as depicted in Figure 1.2 as the distance from the surface increases. The figure shows these two layers and their tangential velocities in the different flow conditions. Notice that the function that describes the change in velocity is non-linear.

1.2.1 Reynolds Number

One very important quantity in fluid mechanics is the non-dimensional Reynolds number which characterizes the interplay between the inviscid and boundary layers. The Reynolds number provides a gauge of the importance of the viscosity compared to the speed of the fluid. The Reynolds number gives us an idea of the state of the fluid.

The equation for the Reynolds number is the following:

$$Re = \frac{\rho V l}{\mu} \tag{1.2}$$

[1] An inviscid region is one where the viscosity effects in the fluid are neglected.

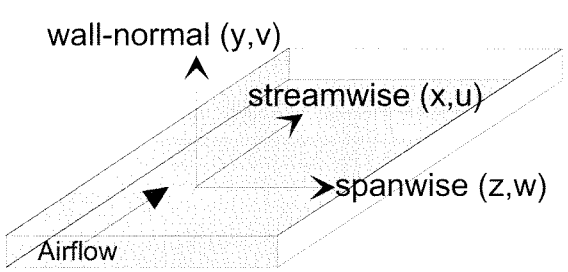

Figure 1.1: The coordinate system depicted in a figure. The stylized wind tunnel is the basis for the coordinate system which also shows the direction of airflow. The coordinates are given as tuples along with the velocity in that direction.

where ρ is the fluid density (equal to $1.18 \frac{\text{kg}}{\text{m}^3}$ for air at 27°C), V is the velocity, l is the characteristic length[2] and μ is the dynamic fluid viscosity (equal to $1.847\ 10^{-5} \frac{\text{kg}}{\text{s m}}$ for air). In Section 1.3 we give the reader an idea of how the Reynolds number varies for our experiments and how the other fluid mechanical quantities vary with the Reynolds number.

1.2.2 Channel Flow

The flow profile we use is called a channel flow. It is generated, for example, in the wind tunnel. Two important properties of a channel flow are:

- The Reynolds number for the flow is calculated from the wind tunnel geometry.
- The wall shear stress is proportional to the pressure gradient.

In the wind tunnel the nature of the channel flow implies $l = \frac{d}{2}$, where d is the channel height. The equation for the Reynolds number in this situation becomes:

$$Re = \frac{U_c\, d}{2\nu} \tag{1.3}$$

[2]The characteristic length is defined according to the flow characteristics. In an flow over a flat plate, l is the boundary layer thickness.

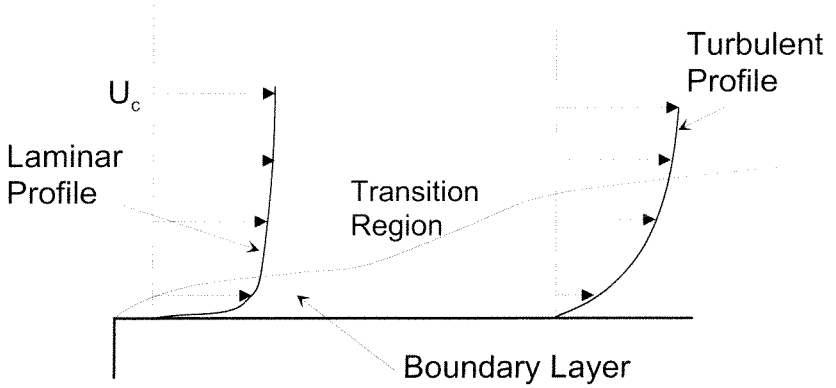

Figure 1.2: The profile of the streamwise velocity (u) in both the linear and turbulent flow plotted over a flat plate.

where ν is the kinematic viscosity equal to $\frac{\mu}{\rho}$

For Reynolds numbers greater than 8000, the flow is considered turbulent. The relationship[3] between the wall shear stress and the pressure gradient allows us to calibrate all of the measurements that we make in the wind tunnel, especially measurements of the wall shear stress.

1.2.3 Law of the Wall

Investigation of the layer closest to the wall[4] leads to a general "law of the wall". This law relates the mean velocity and the components of the velocity fluctuations to the viscous length scale ($\lambda+$). The law describes the nature of the flow from the viscous sub-layer through the buffer layer and into the outer boundary layer.

$$u^+ = f(y^+)$$

The complete description of the function f is generally accepted as:

$$y^+ <= 5 \qquad u^+ = y^+ \tag{1.4}$$

$$5 < y^+ <= 30 \qquad \text{buffer layer profile} \tag{1.5}$$

$$y^+ > 30 \qquad u^+ = \frac{1}{\kappa} \ln y^+ + C \tag{1.6}$$

where $\kappa = 0.40$ is the von Kármán constant and the constant C is approximately 4.5 to 6.0. The buffer layer region is where the viscous and turbulent stresses are about equal making it difficult to derive an equation.

1.3 Range of Variables

In this section we briefly describe the range for values of the fluid mechanics variables. We present the quantities by indexing them by the velocities that are used

[3] See Equation 4.11.

[4] Also known as the laminar sub-layer.

Variable	$U_c = 8\frac{\text{m}}{\text{s}}$	$U_c = 10\frac{\text{m}}{\text{s}}$	$U_c = 12\frac{\text{m}}{\text{s}}$	$U_c = 15\frac{\text{m}}{\text{s}}$	$U_c = 20\frac{\text{m}}{\text{s}}$
Re	6490	8113	9736	12170	16227
U_τ	$.395\frac{\text{m}}{\text{s}}$	$.484\frac{\text{m}}{\text{s}}$	$.572\frac{\text{m}}{\text{s}}$	$.700\frac{\text{m}}{\text{s}}$	$.910\frac{\text{m}}{\text{s}}$
t^*	100μs	66.7μs	47.9μs	31.9μs	18.9μs
y^+	39.6μm	32.3μm	27.4μm	22.3μm	17.2μm

Table 1.3: Table listing the calculated range of the common variables as seen in the experiments as a function of wind tunnel velocity.

in the wind tunnel. Given the geometry of the wind tunnel, the Reynolds number is calculated using Equation 1.3. The other quantities are derived as a function of the Reynolds number. The results are tabulated in Table 1.3.

1.4 Laminar Flow

Low Reynolds numbers are associated with laminar flow. In a laminar flow, the movement of the fluid is said to be smooth and there is very little mixing between the different layers of the fluid. We use the laminar flow regime because it describes a fluid mechanical system with fewer disturbances which means fewer unknowns. For example, generating fixed structures in the laminar flow is useful because the location and nature of the structures generated in the flow field is known since the only ones present are the artificially generated ones.

1.5 Turbulent Flow

High Reynolds numbers are indicative of turbulent flows [3]. The region in between high and low Reynolds numbers is generally partitioned into two subre-

gions, an "unsteady flow" regime and a "soft turbulence" regime. We are interested in studying drag behavior in fully turbulent flows, therefore, we emphasize high Reynolds numbers ($Re > 10^3$).

Turbulence is one of the most vexing problems for fluid mechanicists to understand. Turbulence is exactly what the name has come to imply – non-smooth air flow. Once the flow is determined to be turbulent, the issue is attempting to understand how drag manifests in the flow. Ultimately this understanding raises the bigger question of what methodology allows control of this drag from a practical point of view.

In a turbulent flow, the movement is unsteady due to many different sized eddies. The shear stress at the surface increases and the boundary layer gets thicker as the tangential velocity profile has a higher slope because there is more mixing of the fluid. Another characteristic of turbulence is that flow energy is distributed over many different length scales. The overall nature of a turbulent flow is considered somewhat chaotic. Much research on the chaotic and fractal nature of turbulence over different length scales has yielded positive results in identifying the fractal scale and dimensionality [11]. The length scales in question range from the Kolmogorov length[5] and longer (up to the width of the boundary layer).

Much effort has been spent in understanding the interactions between the fluid in a turbulent boundary layer and the wall [8, 24]. Computational fluid dynamics has been used to study the system [12]. Other researchers use experimental methods to painstakingly study the small-scale structures in the turbulent boundary layer [36].

1.6 Vortex Structures

Vortex structures play an important role in turbulence transport and may cause large skin friction drag. The structures form because disturbances in the bound-

[5]The Kolmogorov length is considered the scale at which the dissipation starts in the energy cascade.

ary layer cause the longitudinal flow to curl into low speed axial flow structures aligned in the direction of the fluid flow. These structures tend to persist for a certain duration before violently bursting into the upper layers of the boundary layer. The interaction of these vortices, which appear randomly in both space and time, with the viscous sub-layer near the surface creates regions of high surface shear stress as depicted in Figure 1.3. These regions are also referred to as streaks.

The shear stress, when integrated over the area of a surface, contributes to the skin friction drag. Attempts to reduce drag by controlling turbulent flows have focused on methods of either preventing the formation or mitigating the strength of these vortices.

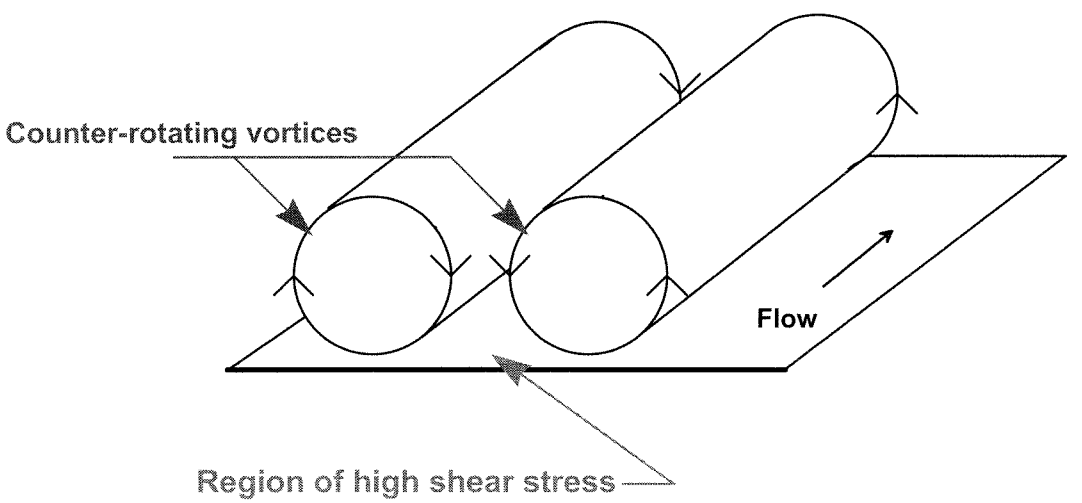

Figure 1.3: A three-dimensional diagram of a vortex pair showing counter-rotating low speed jets of fluid. The region close to the wall between them is where the high speed fluid from above the vortex pair has been carried towards the surface. It is this high speed fluid that causes an increase in the shear stress in that region.

1.6.1 Vortex Statistics

The vortex pairs that are formed in the turbulent air flow of the wind tunnel have microscopic scale.[6] The location of the vortex pairs can be inferred from the measurements of surface shear stress (see Figure 1.3). The spatial statistics of the streaks is depicted in Figure 1.4, Figure 1.5, and Figure 1.6. The different statistics

[6]For that reason they will be referred to as μ-vortices in this thesis.

are reported with respect to the flow velocity to illustrate the microscopic nature of the vortex pairs over a range of low speed flows. In particular, as the flow speeds up, the dimensions of the μ-vortices are reduced.

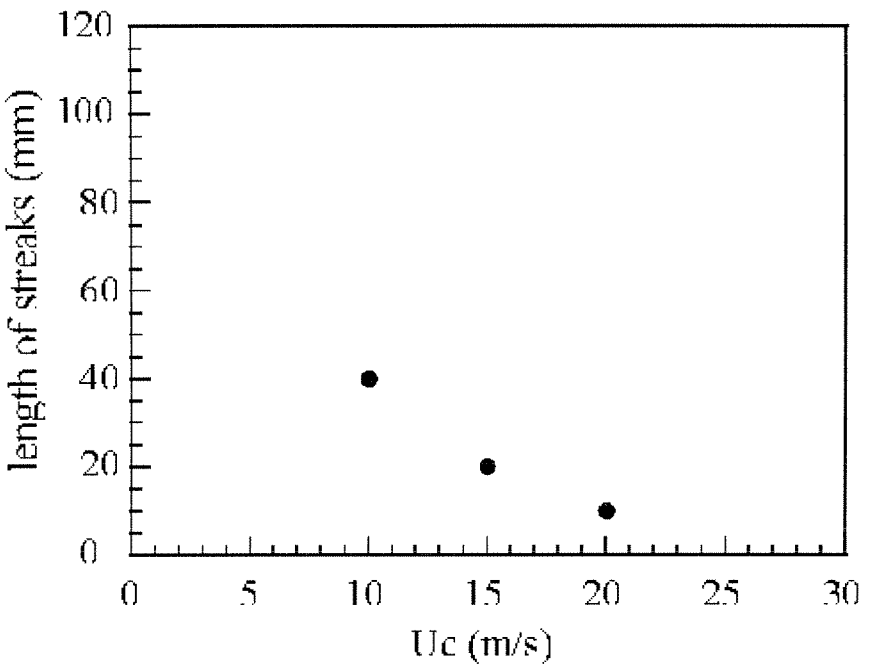

Figure 1.4: The average length of the high shear stress streaks graphed versus wind tunnel centerline velocity (U_c).

1.7 Wind Tunnel

The wind tunnel that is used in these experiments to gather the data is located at University of California at Los Angeles in the laboratory of Professor Chih-Ming Ho. The channel itself is 5 m long with a cross-sectional area of 60 cm x 2.54 cm. Air flow in the channel is generated by an axial blower. The airflow passes through a settling chamber and a 10:1 contraction before entering the experimental area. Maximum velocity in the channel is approximately 24 $\frac{\text{m}}{\text{s}}$ with an rms fluctuation level of about 0.1%. The maximum free-stream velocity used in the experiments presented here is 20 $\frac{\text{m}}{\text{s}}$ ($\approx$ 60 miles per hour).

Two distinct locations in the wind tunnel are used for data collection. The first

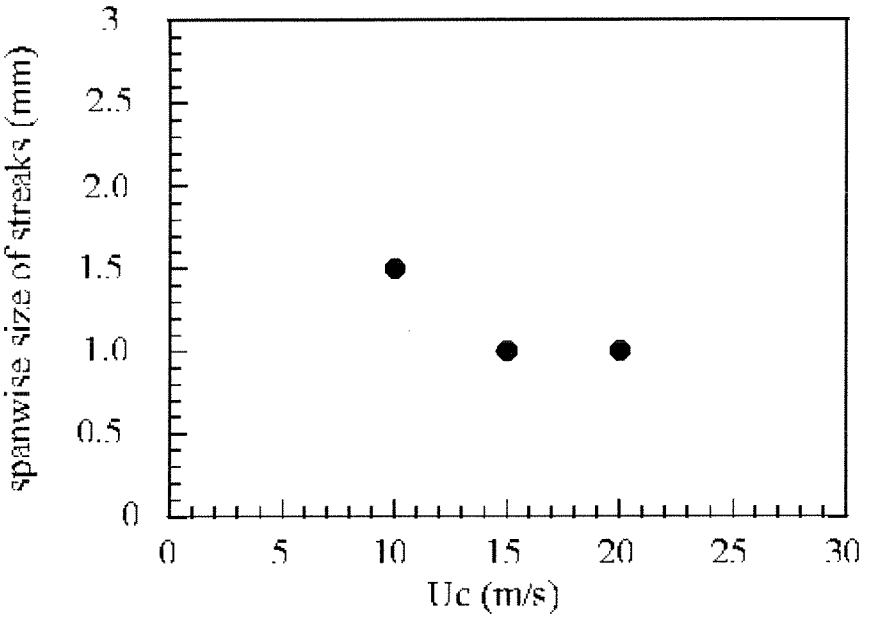

Figure 1.5: The average width of the high shear stress streaks graphed versus wind tunnel centerline velocity (U_c).

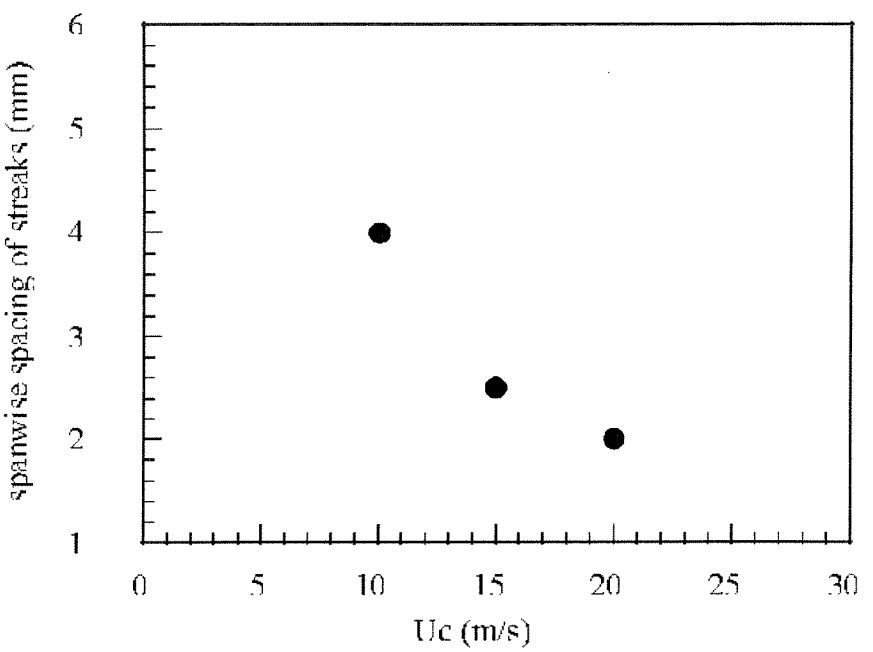

Figure 1.6: The average spacing between high shear stress streaks graphed versus wind tunnel centerline velocity (U_c).

one used for the laminar flow experiments, is located approximately one meter from the beginning of the test channel. The other location is roughly two meters further downstream and it is used for the fully turbulent flow experiments. From the equations of channel flow, a U_c velocity greater than $8\ \frac{\mathrm{m}}{\mathrm{s}}$ produces a flow that is considered fully turbulent at this location [53].

1.8 Spectrum of the Flow

We have recorded the power spectral density of the wall shear stress. Turbulent flows should produce a fairly flat spectral response indicating that the energy is well distributed across all frequencies. We plot various curves which show different frequency ranges in detail (see Figure 1.7, Figure 1.8, Figure 1.9, and Figure 1.10). The different curves on each plot represent different flow velocities (U_c). Notice that the lowest flow velocity ($U_c = 8\frac{\mathrm{m}}{\mathrm{s}}$) in Figure 1.10 has generally more energy around 10 kHz, which implies that at that velocity the flow is not fully turbulent [9].

Figure 1.7 is a spectral response of the system with no airflow present. The data gives us an approximation of the signal floor of our recording system when there is no flow. The next figure (Figure 1.8) demonstrates the power spectrum recorded using our wall-mounted sensors. The curves show that there are no obvious peaks in the spectra. This flatness (up to about 10 kHz) means the flow can be considered somewhat random[7]. Figure 1.9 and Figure 1.10 are plots which emphasize low and high frequencies respectively.

1.9 Intermittency Functions

A number of researchers have tried to indicate the presence of a turbulent region by defining intermittency functions. The use of the intermittency functions is generally confined to the transition region between laminar and turbulent flow regimes.

[7] A truly random flow would have a flat spectrum across frequencies of interest.

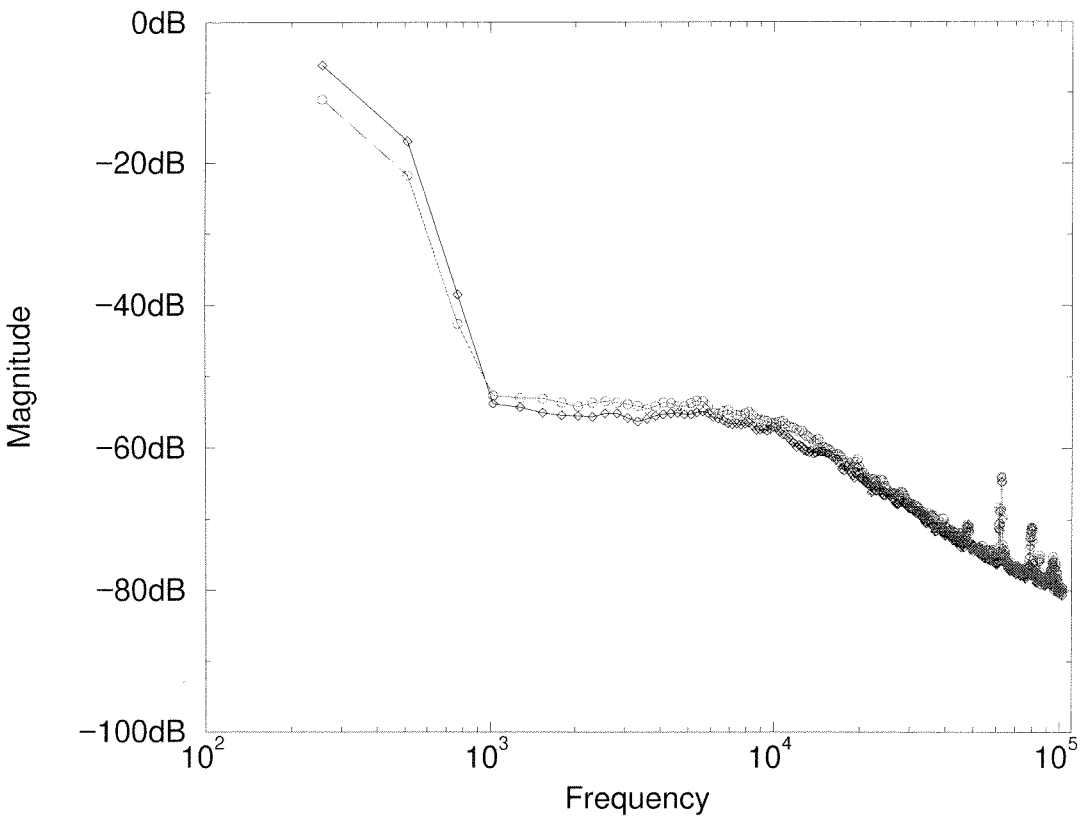

Figure 1.7: Spectral measurement of the sensor system in the wind tunnel with zero velocity. This plot shows that there are no salient signals introduced by the measurement system in either identically measured curve.

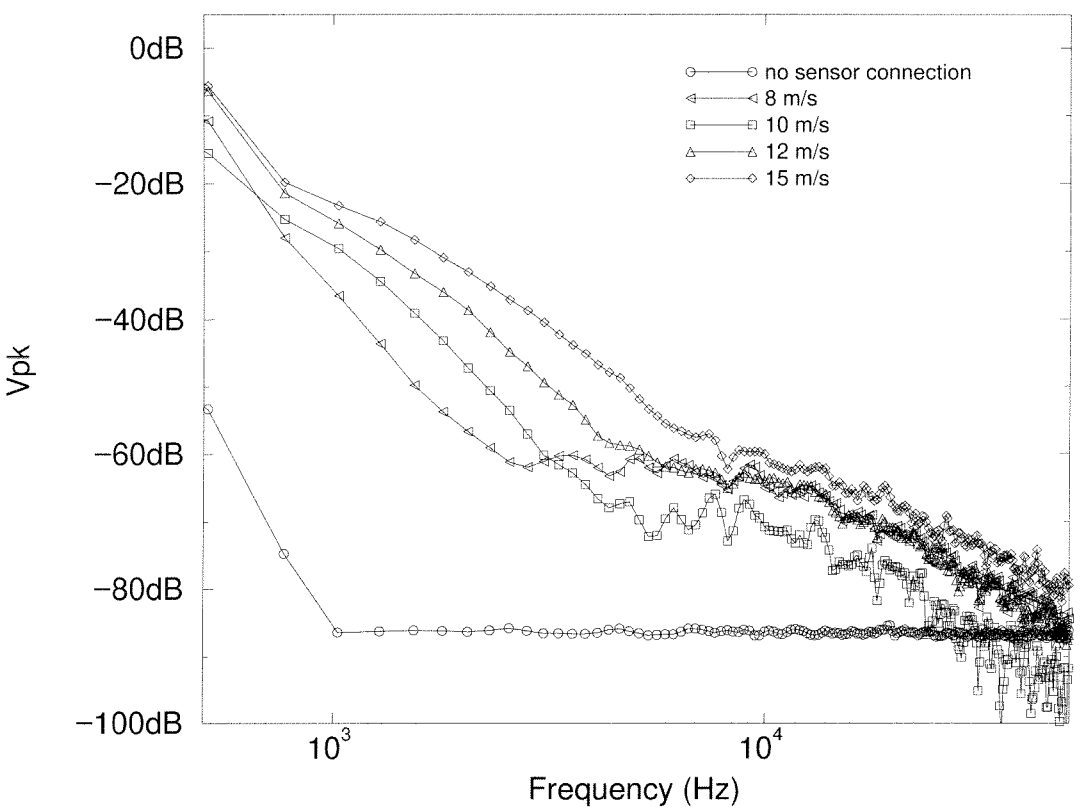

Figure 1.8: Frequency response of our system in the wind tunnel at various speeds. The data has been normalized by the curve shown in Figure 1.7. The curve labeled *no connection* represents the noise floor for the system with no sensor attached.

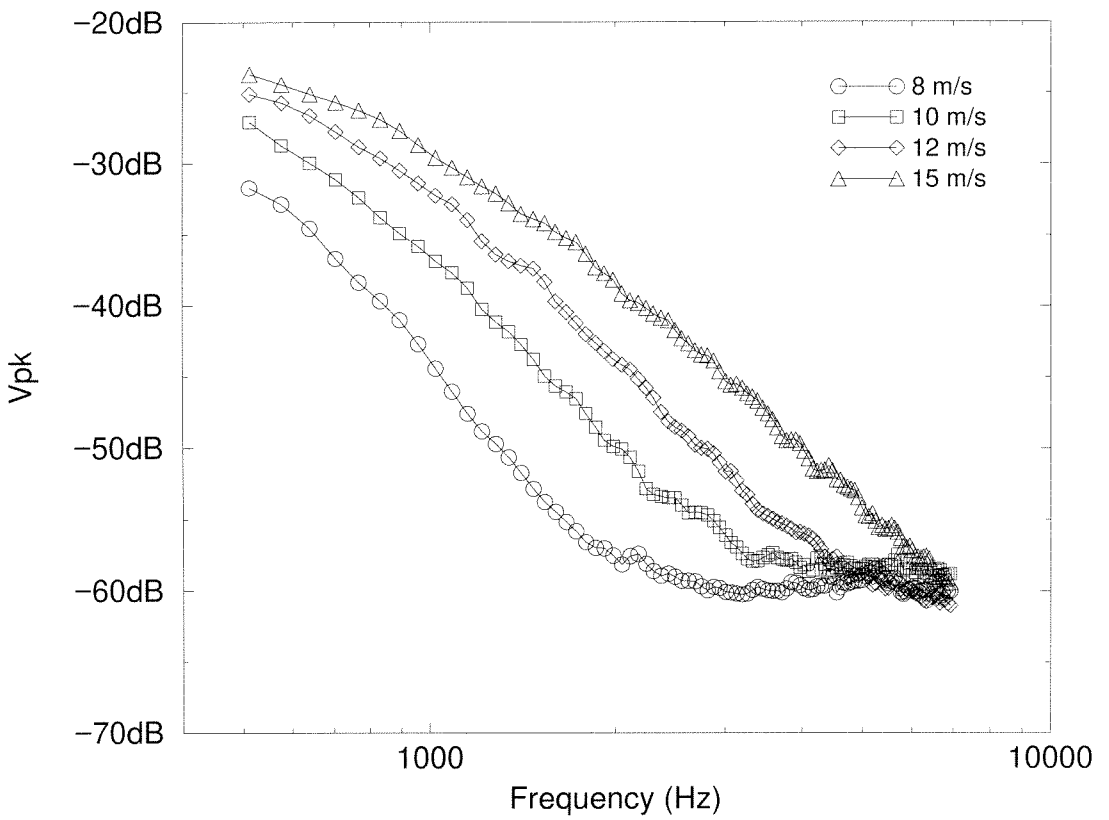

Figure 1.9: Frequency response of our system in the wind tunnel at various speeds with emphasis on lower frequencies.

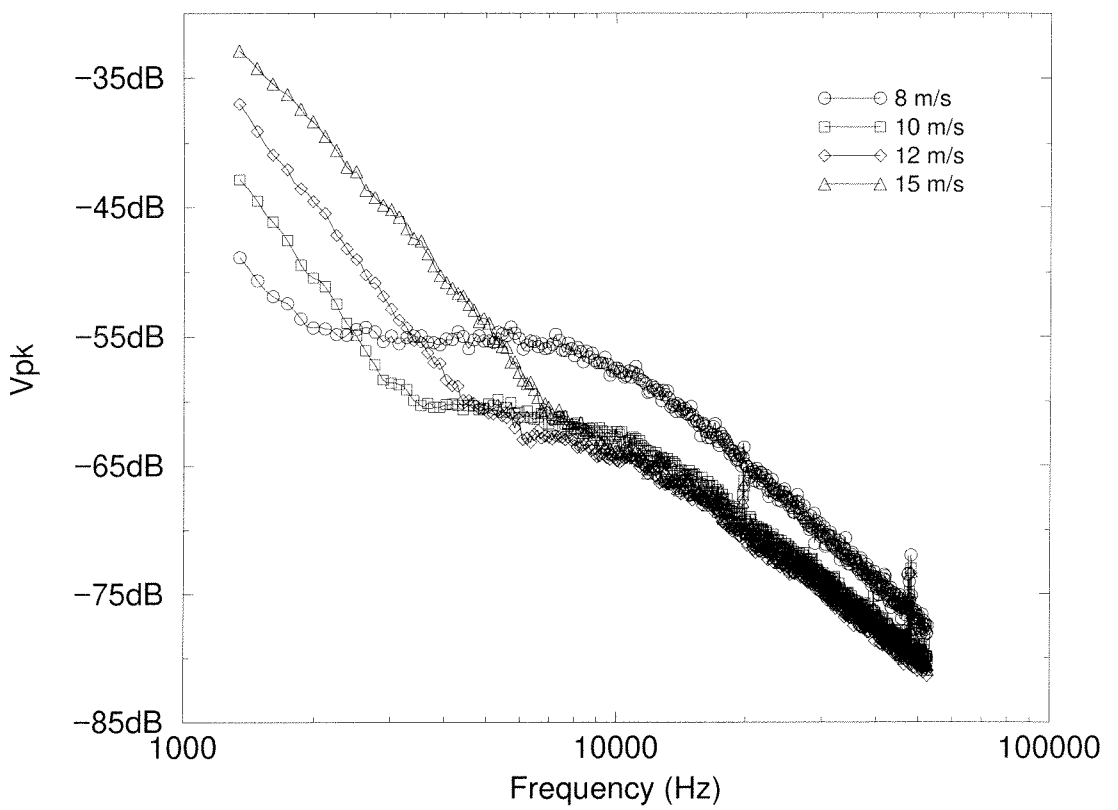

Figure 1.10: Frequency response of our system in the wind tunnel at various speeds with emphasis on higher frequencies.

Author	Intermittency Function
Townsend (1949)	$\lvert u\rvert, \lvert\frac{\partial u}{\partial t}\rvert$
Corris and Kistler (1955)	$\lvert u\rvert, \lvert\frac{\partial u}{\partial t}\rvert$
Keskestad (1965)	u^2
Gartshore (1966)	$\lvert\frac{\partial u}{\partial t}\rvert$
Fiedler and Head (1966)	$\lvert\frac{\partial u}{\partial t}\rvert$
Kaplan and Laufer (1968)	$\frac{\partial u}{\partial t} - \{\frac{\partial u}{\partial t}\}^2$
Wygnanski and Fiedler (1970)	$(\frac{\partial u}{\partial t})^2 + (\frac{\partial^2 u}{\partial t^2})^2$
Kovasznay *et al.* (1970)	$\lvert\frac{\partial^2 u}{\partial t \partial y}\rvert$
Antonia and Bradshaw (1971)	$(\frac{\partial u}{\partial t})^2$
Antonia (1972)	$(\frac{\partial uv}{\partial t})^2$
Thomas (1973)	$\lvert\frac{\partial u}{\partial t}\rvert$filtered
Bradshaw and Murlis (1974)	$\lvert\frac{\partial uv}{\partial t}\rvert$or$\lvert\frac{\partial^2 uv}{\partial t^2}\rvert$

Table 1.4: Table of intermittency functions used by fluid mechanicists to detect turbulent regions. Reprinted from [9].

These intermittency functions are tabulated in Table 1.4. The functions generally involve the time derivative of stream-wise velocity ($\frac{\partial u}{\partial t}$). It should be noted that most of these intermittency functions are formulated using the output signal of true hot-wire probes, hence the availability of $\frac{\partial u}{\partial t}$ values. If only surface shear stress data is available, then the measurement of $\frac{\partial u}{\partial y}$ is more appropriate for incorporation into an intermittency detector function. The intermittency functions will form a basis for the type of functions that we might wish to implement.

Chapter 2 The Shark

In trying to solve many complex problems in the world, one can often find inspiration by observing how certain animals have evolved through natural processes to address the problem in different ways. Looking at animals for innovative ideas about drag reduction is not a new concept [10]. There are many examples of how different species use a variety of methods to reduce drag. Porpoises, dolphins, and sharks are examples of animals having potentially drag reducing skin. For instance, the compliant skin of the dolphin and porpoise is conjectured to be used to dampen out turbulence on its surface. The purported evidence of this function is the discrepancy between the top speeds the dolphin and porpoise are predicted to attain versus what they actually achieve. Current research tends to discount this apparent gain because of faulty top speed prediction.

We look to the shark as a biological inspiration to the problem of drag reduction. In particular, fast sharks serve as a good biological model because they are a highly evolved predator with at least a 350 million year old history of the genus [17]. The predicted Reynolds number for the flow over a shark's skin can be as high as 10^6 to 10^7 [6].

Since the Triassic period[1], sharks have been evolving as ocean-going creatures. Very little is known about the shark forerunners derived from the neoselachians, so evolutionary biologists assume that the shark forerunners had a very similar ecology as modern day sharks. There are several distinct types of modern day sharks:

- reef dwellers (durophagous[2] type)
- demersal sharks (on soft bottoms)

[1]The geologic period ranging from approximately 245 to approximately 208 million years ago.
[2]Possessing a hard shell.

- fast off-shore predators
- mid-water and deep-water sharks usually bioluminescent
- plankton feeders[3]

2.1 Off-Shore Sharks

Of the many families of modern day sharks, we are most interested in the ones which are categorized as fast off-shore predators. They seem to have evolved differently from the other families of either fresh-water or coastal sharks. Examples of specific families of fast off-shore predators are as follows:

- *Orectolobidae*
- *Triakidae*
- *Carcharhinidae*
- *Sphyrnidae*

These sharks inhabit the oceans at a depth from the surface down to 200 meters. Man rarely comes into contact with them because of their relative small numbers and adjoint habitation areas. These sharks (for example, hammerhead sharks) can swim up to 20 $\frac{\text{m}}{\text{s}}$ (72 km per hr) in deep water. The exact nature of the physiology of these sharks remains a mystery because they are difficult to study as the ocean setting is hard to replicate in a controlled environment such as a marine laboratory.

2.2 Shark Skin (Dermal Denticles)

Marine biologists have studied the shark's skin for a number of years. The unavailability of fresh samples which showed no degradation of the fine structure often

[3]e.g., whale sharks, devil rays, and basking sharks.

led the biologists to merely label the skin as abrasive. The shark skin is covered by scales called dermal denticles (Figure 2.1). The basic structure to the scale is a mast which extends from the skin attached to a plate which lies parallel to the shark's body. The shape of this plate varies depending on the shark and the scale location. The plates on the sharks of interest have ridges on them aligned in the direction of the water flow. It is the microscopic nature of these ridges which is exciting. Both the microscopic ridge height and ridge to ridge spacing vary depending on the shark species. The ridge spacing for fast off-shore sharks ranges from 35 μm to 105 μm (e.g., Figure 2.2).

Only since the mid 1970's has scanning electron microscopy imaging been applied to the study of the complex microstructures on dermal denticles. Researchers have also found interesting drag-reduction properties resulting from the microscopic structure of the dermal denticles [48, 47, 46, 16]. Researchers have also recently examined the role of the skin in shark locomotion as being similar to that of the dolphin and porpoise, i.e., a compliant surface [55].

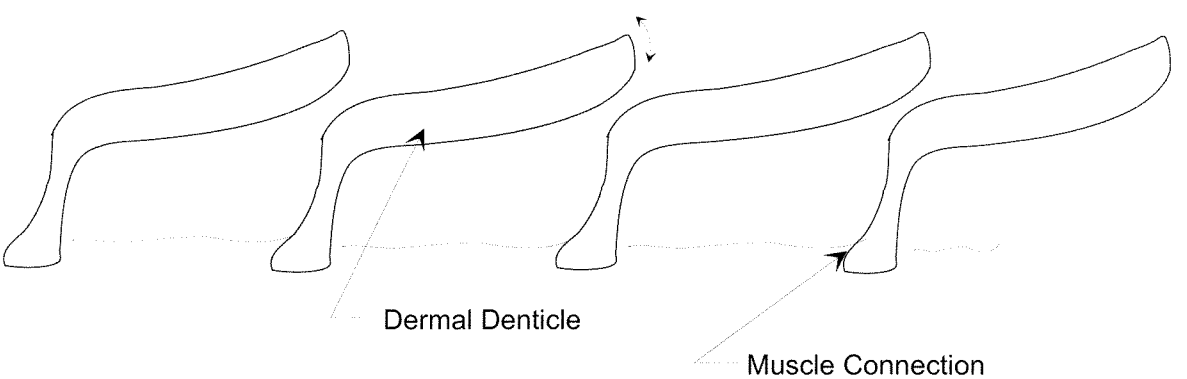

Figure 2.1: A side perspective drawing of a shark's dermal denticles showing how they overlap each other and protrude from the shark's skin.

Like every highly evolved system in nature, the scales of the sharks are thought to serve several complementary purposes. Among these functions are:

- scratch and abrasion protection from the environment – rocks, coral, etc.
- parasite and predator protection – microorganisms as well as smaller fishes

- drag reduction – the most controversial yet interesting function

The natural argument using Darwinian evolution would lead one to believe that the structure of these scales assists the shark's movement making the shark a more efficient predator. This assistance in movement may be indicative of some method of drag reduction. Research into passive drag reduction has helped the understanding of the nature of the structure of the shark skin [6] (Section 3.2.2).

2.3 Pit Organs

Pit organs are found over the entire body of off-shore predator sharks. They are located in microscopic channels under specially modified dermal denticles (see Figure 2.3). The pit organ is essentially a small, several-celled sensor known as a free neuromast. A typical pit organ has a diameter of about 110 μm [52]. Their exact function is not known to the shark biologists. Some researchers speculate that they are used by the shark to detect displacement and location of a water disturbance [52]. Other researchers relegate them to the simple role of salinity detectors [33]. These researchers have also thought the pit organs act as both mechanical as well as chemical sensors[4] [32].

A more recent theory proposed by Reif claims that the pit organs function as near-field detectors of the water flow [47]. The evidence for this hypothesis is as follows:

- Pit organs are located over the entire body, especially on the backside of the shark.
- Positive correlation between the maximum speed and number of pit organs on the shark.
- The specific pattern of pit organs corresponds to locations where turbulence might be expected.

[4] A complex role which is carried out by a mechanoreceptive sensor.

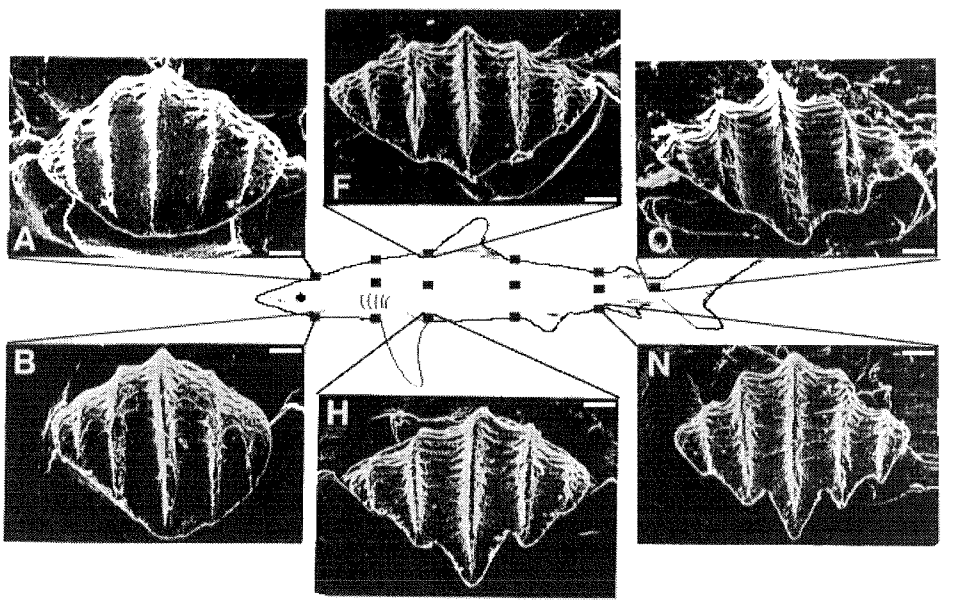

Figure 2.2: The dermal denticles from various parts of a sharks' skin. The shark is a *Carcharhinus falciformis*. The figure is reproduced from [46].

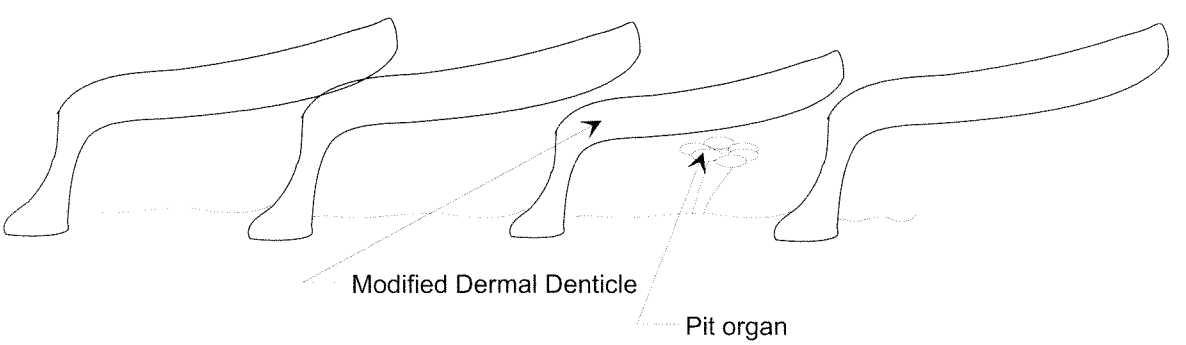

Figure 2.3: A side perspective drawing of a shark's dermal denticles and a single pit organ. The drawing shows how the pit organ sits underneath a modified denticle which is generally smaller than a normal denticle.

species	max m/s	pit organs
Ginglymostoma cirratum	10	150-300
Carcharhinus miliberti	13	550
Isurus oxyrhinchus	15	800
fast *carcharhinids*	20	700-1300
hammerheads	20	700-1300

Table 2.1: Table of the relative speeds of different sharks along with the pit organ counts which shows a positive correlation between the two quantities. The data is extracted from several different studies [47].

2.4 Active Control

The question of active control of shark skin is very much a speculative one. Because most sharks are either severely damaged or dead by the time marine biologists have an opportunity to examine them, the shark pathology is woefully incomplete. However, biologists hypothesize that sharks can actively move their denticles. The indirect evidence of this control is twofold. First, the denticles connect to muscles underneath the shark's skin. Second, the total number of pit organs and their placement on a shark's body positively correlates with the speed of the species (Table 2.1). This correlation is consistent with the theory that for active control the shark would need many sensors to relay the current conditions over its body.

However, questions do remain about sharks utilizing active control. The conclusion that one can draw from this biological example is that it may be beneficial to employ actively controlled microscopic structures to reduce drag.

2.4.1 Blowing and Suction on the Shark Skin Surface

Bechert has proposed a novel hypothesis about another mechanism that may aid drag reduction on a sharks' skin [6]. He proposes that there is a secondary flow that exists underneath the dermal denticles that can either blow fluid up into the main flow or suck fluid from the main flow (Figure 2.4). This "streak cancellation"

mechanism is likened to a surface jet of fluid at an acute angle to the sharks skin. This surface jet impinges at locations between the edges of the dermal denticles. The evidence that he points to is the apparent existence of fluid channels *beneath* the dermal denticles, the particular shape of the denticles, and how they fit together.

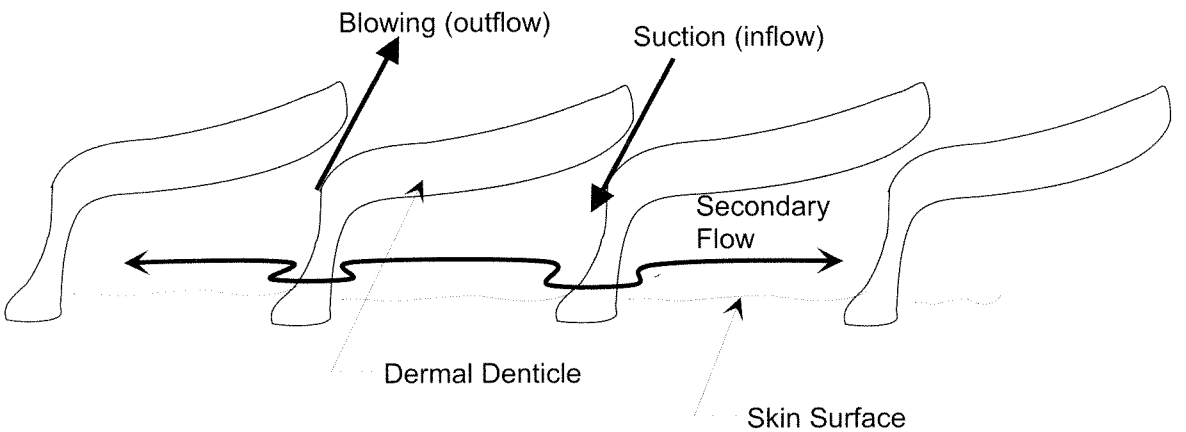

Figure 2.4: A drawing of Bechert's hypothesis for the surface jet on a shark's skin. The drawing shows how the fluid might pass between the dermal denticles producing both blowing and suction.

Chapter 3 Drag Control Approaches

Researchers in fluid mechanics have tried many approaches to the challenge of drag reduction both theoretically and experimentally [22]. Because of the inherent complexity of the problem, the approaches have not met with a great deal of success especially in the commercial arena. This chapter describes some of the attempted solutions to the problem and why our alternative approach may be fruitful.

3.1 Numerical Approaches to Drag Reduction

One tool that can help fluid mechanicists explore the question of how to achieve drag reduction is computational fluid dynamics (CFD) simulation. The simulation environment provides approximate solutions to the basic equations in fluid mechanics, i.e., the Navier-Stokes equations as well as conservation of mass, vorticity relationships, etc. With the advent of powerful computers and accurate numerical methods, the CFD field has been used extensively to study phenomena that would be otherwise difficult to reproduce in the laboratory[1].

Typical simulation resolutions used in our research is 32 spanwise coordinate points by 32 streamwise coordinate points by 32 wall-normal coordinate points. The CFD simulations produce structures and behavior that agrees well with experimentally observed phenomena [35]. The accuracy and reproducibility of the CFD simulations provides a good platform for testing drag reduction algorithms.

[1] e.g., a wind or water tunnel.

3.1.1 CFD Drag Reduction Methodology

CFD simulations carried out by other researchers demonstrate that suppressing the interaction between stream-wise vortices and the wall achieves significant drag reduction (on the order of about 25%) [13, 14]. These experiments incorporate active feedback control to achieve this drag reduction performance. The control scheme involves imposing a boundary condition in the simulation which mimics blowing and suction at the wall according to the normal component of the near-wall velocity field ($-v(y^+ = 10)$). Figure 3.1 demonstrates the control algorithm. The rationale behind this simple control law is to push the turbulent structures responsible for the high shear stress away from the wall. This strategy forms the basis for the system that we wish to build.

The drawback of the control law as described in this section, however, is that it is physically unrealizable. Simply stated, there are no viable techniques that can measure the near-wall normal velocity of a fluid without disturbing the fluid, and hence increasing drag. Obviously, some modification of the control algorithm is necessary for it to be useful in a real-world system.

Another practical issue with the CFD simulation is that both the sensing and actuation have greater precision than what is achievable in a real system. In other words, the sensors have infinite dynamic range[2]. With the imposition of the condition of no net mass transfer at the wall[3], the actuators are modeled as simple wall velocity sources or sinks with infinite dynamic range and no delay in actuation.

The actuator is incorporated into the simulation as boundary condition (i.e., $v(y^+ = 0)$ is equal to the imposed velocity at the wall), at each wall grid point in the computational domain. The real devices that we can build are significantly different, however, the simulations would become too expensive and slow if one tried to model the actuator in a form that approached the actual physical device.

[2] At least to the floating point precision of the machine being used for the simulation.

[3] In other words, the total mass flowing through the wall boundary has to sum to zero at all times.

3.1.2 Linearized Control Approach

One common approach in the attempt to find tractable, closed-form solutions to the equations of fluid flow with disturbances is to apply classical control system theory to find a set of suitable control laws of a reduced-dimensionality system [31, 1]. Given that the Navier-Stokes equations are a set of non-linear partial differential equations, a straightforward application of a linear control system is not feasible. Linearizing the equations about an operating point is one way to develop a reduced complexity model. Other simplifications can be made. For example, using a 2-D Poiseuille flow[4] profile with an artificial stimulus providing some randomness instead of a fully 3-D turbulent model is one such simplification. Even with these simplifications the linear model complexity approaches hundreds of poles and zeroes in order to model the flow with a reasonable level of detail. Obviously, one major drawback to this approach is the inability to approximate the full stochastic nature of real turbulent flow.

The approximate controllers derived using this methodology must also obey the constraints present in the CFD code. With any CFD approach, the sensing and actuation that can be simulated are artificial at best. Because the real sensors and actuators have many limitations that cannot be modeled easily in a simulation, the

[4]Poiseuille flow is what results when a flow is driven by an externally imposed pressure gradient through two stationary walls.

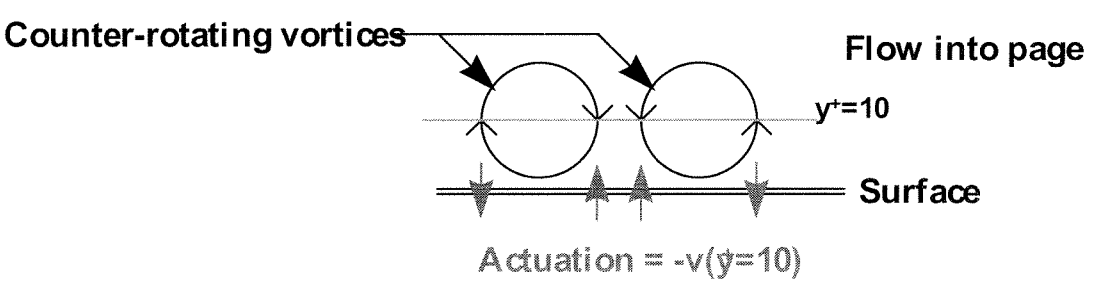

Figure 3.1: Depiction of an actuation law that is used in a computational fluid dynamics simulation which reduces drag by $\approx 25\%$ [35]. The simulation is conducted using a fully turbulent flow. The control law uses measurements from a plane located at $y^+ = 10$ within the fluid to achieve a good drag reduction value.

control results cannot be directly applied to a real system. However, the simulations will hopefully enumerate basic principles that will provide guidance in real system design.

3.1.3 Neural Network Models

In other work assisted by the author, neural networks were applied to the problem of drag reduction in CFD simulations. The focus of that research was to determine if a pattern can be detected in some ancillary measurements that allows identification of high shear stress regions.

Other researchers have used neural networks to develop a control scheme that attempts to minimize wall shear stress [27]. The controller was incorporated into a pseudo-turbulent 2-D flow simulation with periodic disturbances. However, such a simulation is fundamentally different from a fully turbulent 3-D simulation.

In our work, we have used a simulation of a fully turbulent 3-D flow. We extract the surface measurements of shear stress $\frac{\partial u}{\partial y}$ and $\frac{\partial w}{\partial y}$ for use as inputs training data. Our target data is the corresponding velocity to be imposed at the surface. This work attempts to approximate the control law discussed in Section 3.1.1 by correlating surface measurable quantities with the desired wall velocity vector.

The basic network architecture employed is one called a shared weight template. This type of network is basically a trainable local feature detector which is very efficient in its use of storage. Furthermore it inherently assumes translational invariance, i.e., looking for a similar pattern anywhere along the wall surface.

Neural Network Topology

In applying a neural network to the problem, we are building a pattern detector. A generic architecture is depicted in Figure 3.2.

As inputs to the network, we provide quantities that we will be able to measure at the surface of the wall. These quantities are the shear stress quantities (both stream-wise and spanwise) $\frac{\partial u}{\partial y}$ and $\frac{\partial w}{\partial y}$. A plot of the sample data is in Figure 3.4.

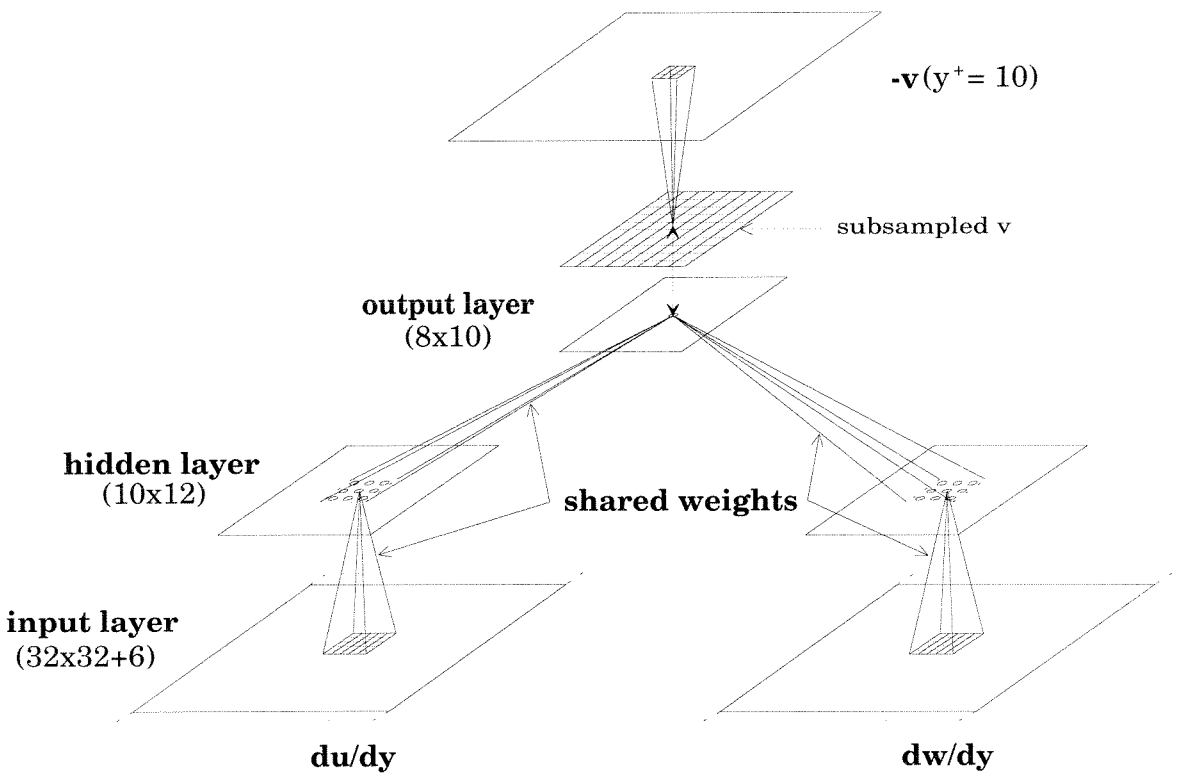

Figure 3.2: Generic network topology used as a feature detector to predict the wall velocity necessary to reduce overall drag. For the most part, the weights are used in accomplishing the pattern detection.

For a particular set of on-line experiments, the topology of the network consisted of the inputs from a one by n slice of the $\frac{\partial w}{\partial y}$ data (Figure 3.3). The prediction with this network is reasonably good considering how few weights are actually used [2, 37, 38].

Network Weight Pattern

Examining the weights learned by a network that shows reasonable drag reduction performance, we note some interesting features. The most important one is that the function exhibited by the weight values is an approximation to the span-wise derivative (see Figure 3.5). In other words, the learned pattern detector relies on large spanwise differences to accurately gauge when and where to actuate in order to reduce the wall shear stress. Also, for networks with more than five spanwise units, the function approximated is very similar [2]. The spanwise derivative is a function that can be implemented in a hardware system.

NN Issues

One concern with these results is the CFD simulations generally use a Reynolds number that is significantly less than what we encounter in the wind tunnel due to the expense of supercomputer time. Thus it is hard to expect that direct application of information learned in the CFD simulation to the control circuits in the wind tunnel will give similar results. Also like the linearized control approach, the real limitations of the sensors and actuators are not modeled.

However, the general principles learned may still be valid. For instance, the region of high shear stress corresponding to the location of μ-vortices is indicated in both the CFD simulation and data recorded from the wind tunnel (see Section 6.3.1).

3.2 Experimental Drag Reduction Techniques

3.2.1 Large Eddy Breakup Devices

One very basic method for trying to reduce the wall shear stress is to introduce a device into the flow boundary layer which attempts to "calm" the fluid flowing past it [56]. One such device is called a large eddy breakup device (LEBU). The basic problem with a LEBU is that no one to date has been able to demonstrate a LEBU design that reduces more drag downstream than the placement of such a device creates within the flow. In other words, there have been no positive net gain experiments.

3.2.2 Riblets – Experimental and CFD Studies

Among the passive techniques for drag reduction, riblets are the most prominent. A riblet covered surface is one that has small (even microscopic) grooves patterned on it in the direction of the fluid flow. Fluid mechanicists have discovered that such a surface leads to a reduction in drag both in the CFD simulation as well as the wind tunnel [35, 54, 5]. Simulation results show about a seven percent reduction in total drag at a fixed velocity [35]. It is also clear from fluid mechanical simulations of the shark's dermal denticle that a ridge-like shape aligned to the fluid flow can reduce the drag.

Researchers have also examined the effect of riblets on airflow over a model aircraft [54, 5]. Of special interest is the influence riblets have on a turbulent boundary layer. The principle behind riblets are to modify the behavior of the vortex pairs. This technique has been verified by experimental research, generally finding a reduction in the drag of around 7 - 8% [4, 51].

The reason for this drag reduction is the nature of the grooves reduces the contact area of the μ-vortices with the surface (see Figure 3.6). In effect, the μ-vortices get "pushed" off of the surface. Were the μ-vortices to attempt to come closer to the wall, the geometry effectively slows down the vortex pair. This minimization

of contactable surface over which the high shear stress is distributed leads to a reduction in drag. Using this concept to build a system except substituting an active expulsion of high shear stress for the passive mechanism, would provide an effective answer to the dilemma of drag reduction.

There are several problems in using riblets in a commercial setting. First, since the groove spacing is fixed, the riblet surface will work optimally only for a narrow range of velocities where the spatial statistics of the μ-vortices matches the predefined riblet spacing. Second, the riblet surface increases maintenance costs because the grooves must be kept clean in order for the riblets to be effective.

3.3 Other Experimental Drag-Reduction Attempts

In addition to passive means for reducing the turbulent drag, a number of researchers have also experimented with other, more sophisticated, active control means to achieve the same goal.

3.3.1 Selective Suction

Researchers have attempted to fashion control systems using mechanical devices to reduce the effect of streamwise vortices. In one method suction is utilized by the researchers in an attempt to control the turbulent events (of vortex formation and breakup) in an attempt to reduce the drag [25]. Other researchers similarly use selective suction to control the emulated streamwise vortices [43]. The control follows a simple rule that is described as continuous suction in selective points with the control imposing a small wall normal velocity. The important point to note with these experiments is that with the application of small amounts of suction, positive results (up to 5% drag reduction) can be obtained in controlling the flow. Suction velocities as low as $6\,10^{-4}\,U_c$ have been reported as being efficacious.

The basic issue with these attempts to use suction and blowing to control the drag is one of scale. If the blowing and suction systems can be built on a micro-

scopic scale with the appropriate control, then this approach has the scalability to be useful in drag reduction. At this point the experiments have been conducted at a scale much larger than practical for our purposes. Another concern is the experiments conducted so far have used artificial methods to generate the vortex pairs in laminar flow, hence there may be problems using similar techniques in a fully turbulent flow.

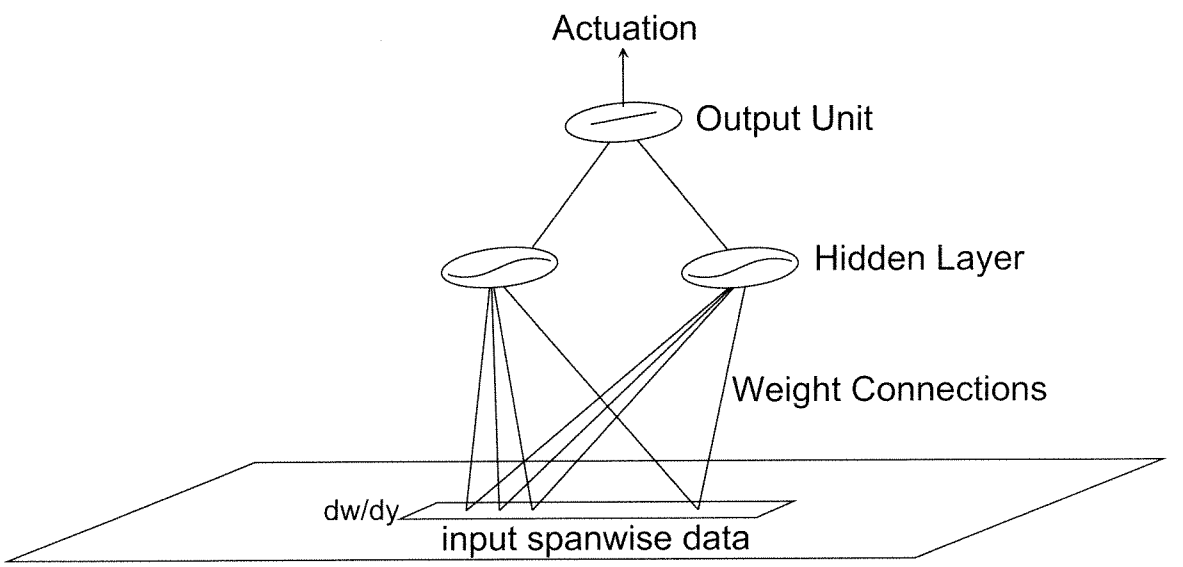

Figure 3.3: Specific network topology used in the flow control simulation that predicts the wall velocity necessary to reduce overall drag by approximately 20%.

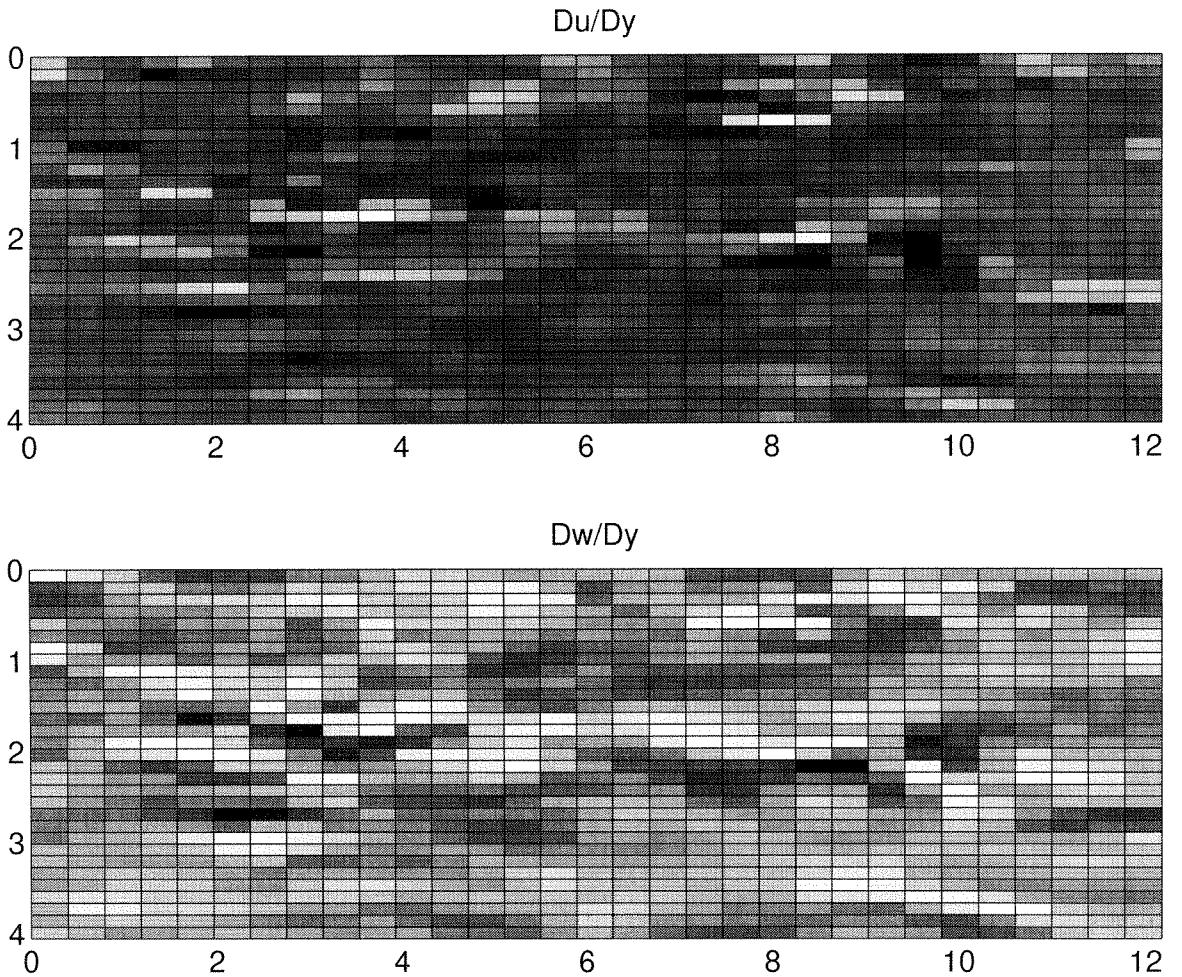

Figure 3.4: Two-dimensional plots of different surface shear stress patterns (both $\frac{\partial u}{\partial y}$ and $\frac{\partial w}{\partial y}$) used as training inputs to a neural network. The neural network trains in order to find a correspondence between input patterns and the predicted wall actuation with the goal of constructing a network that reduces drag. The data is plotted with the z dimension on the vertical axis and the x dimension on the horizontal axis. The data ranges from low values drawn in black to high values drawn in white.

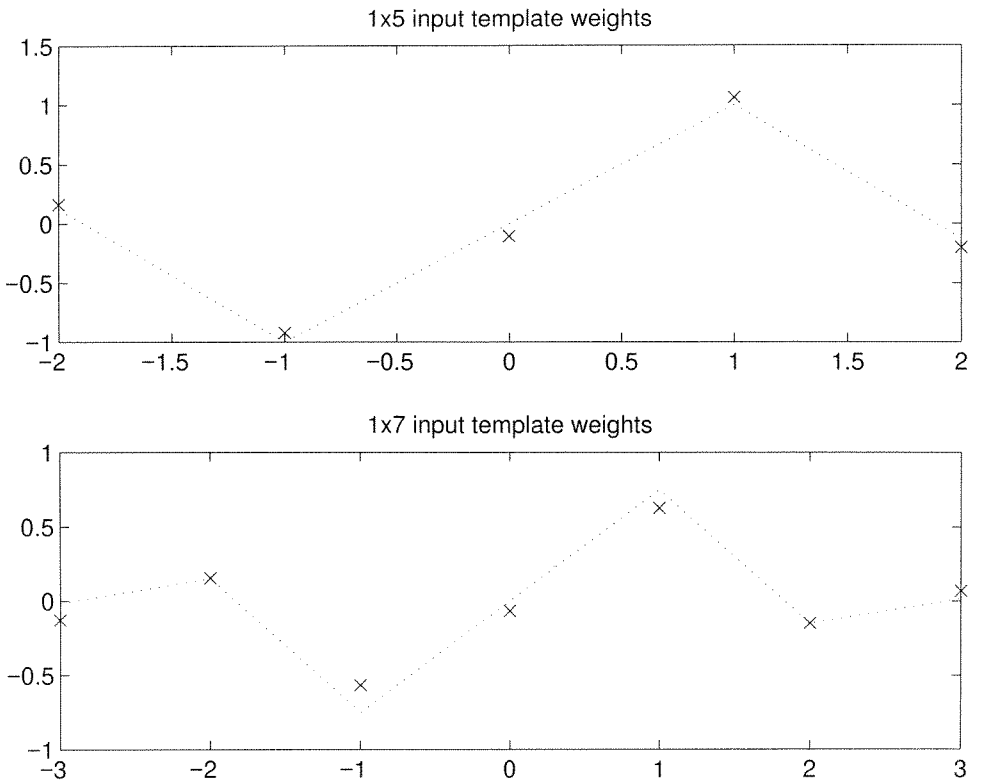

Figure 3.5: The weight magnitudes for a learned NN controller plotted versus their spanwise location in the weight template. Note that the weights mimic the curve constructed from the Taylor coefficients of the corresponding approximation to the span-wise derivative (dotted line).

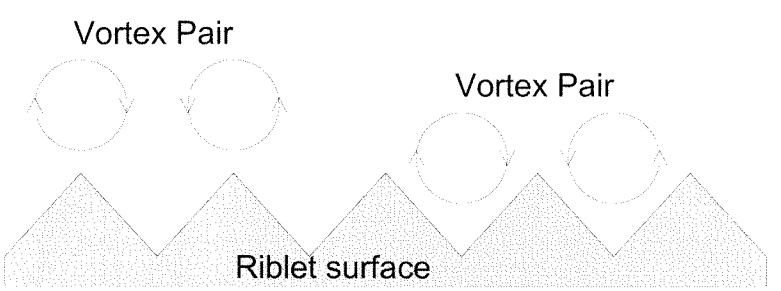

Figure 3.6: Riblet surface depicted interacting with μ-vortex pairs. The spacing of the riblets is important as it should match the μ-vortex dimensions.

Part II

System

Chapter 4 Micro Electro Mechanical Systems

In order to tackle the problem of turbulent drag reduction, our goal is to build an integrated system. This system is comprised of sensors, actuators, and signal processing circuitry between them. Figure 4.1 is a simplified diagram of the proposed system.

Micro-electromechanical systems (MEMS) are used to build the sensors and actuators for the simple reason that the devices can be built with sensing ability and actuation range comparable to the spatial scales of the naturally occurring phenomena, the μ-vortices. This is an important point to note. A fair amount of drag reduction research has relied on "large" mechanical devices acting on fluids with artificially large length scales in attempts to address the problem. It is clear that if the actuator is an order of magnitude larger than the size of the μ-vortices present in the flow, then the problem of spatially-localized control is difficult. Likewise if the sensor cannot resolve the structures either spatially or temporally, the ability to detect the transitory regions of high and low shear stress will be impaired [50, 49]. Furthermore if the sensor interferes with the flow and corrupts the surface shear stress measurements, the problem becomes much harder to solve [41].

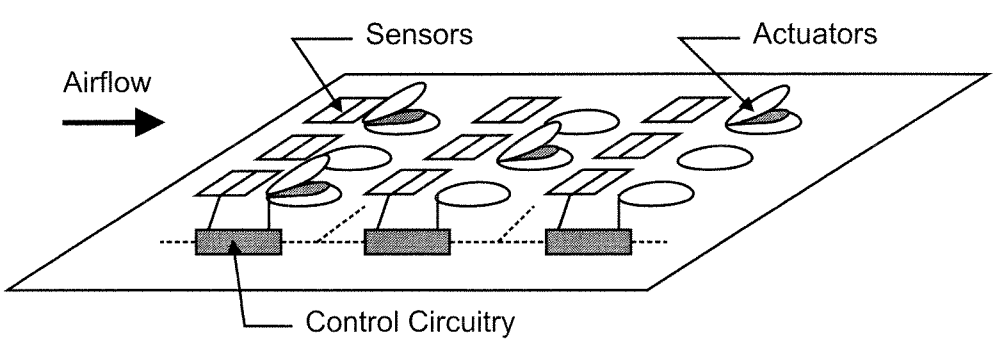

Figure 4.1: A general picture of the desired system showing sensors, actuators and control circuitry colocated on the same substrate.

4.1 Micro-ElectroMechanical Technology

The recent advent of micromachining technology has spurred research into applications to systems with small mechanical scales. In particular, the technology allows construction of shear stress sensors and actuators on the same microscopic scale as the physical phenomena we are trying to control.

The basis for this technology is the materials and processes used in standard CMOS fabrication. Hence, the biggest advantage of this technology is the ability to exploit years of research into fabrication techniques and materials. This technology provides innovative microscopic structures that can be built to perform unique tasks. Examples of some devices include μ-pressure sensors, accelerometers and μ-mirror based imagers [57].

The thrust of this chapter is to give the reader an introduction to the specific MEMS devices that will be used with VLSI circuits to build our integrated system. In particular, we describe some of the properties of the devices, especially those relevant to the circuit design. The μ-sensors and μ-actuators, constructed at Caltech by researchers in Prof. Y.-C. Tai's group, employ both surface and bulk silicon micromachining technology [39, 40, 28, 29, 30]. Further detail for both the μ-sensors and μ-actuators including the fabrication aspects of the components can be found elsewhere [40, 39, 28, 29, 30].

4.2 Shear Stress Microsensor

The μ-sensor (Figure 4.2) facilitates measurement of the heat transfer between a polysilicon element on the silicon surface and the air flowing over it. The electrical properties of the wire change in a detectable manner as this heat convects from the electrically heated wire to the fluid flow.

According to Ohm's law, as a voltage is applied to a resistor the current flow in the resistor is inversely proportional to the resistance of the element. The resistor in our sensor is a length of polysilicon wire. The fabrication procedure for polysilicon

deposition allows the resistance to be specified by changing the doping level and several other parameters. The current that flows through the polysilicon resistor causes power to be generated in the resistor which is dissipated in the form of heat. This additional heat causes an increase in the operating temperature of the wire. The important physical property of polysilicon for our application is that the resistance is a function of the operating temperature. This property allows the polysilicon to function as a sensor by heating up the wire with a known amount of power and then measuring the effect of the fluid in the environment on the sensor's temperature[1]. In our case, if the fluid cools the sensor, the sensor resistance will decrease. A sensor is characterized by a cold resistance R_{cold} at a temperature of T_0.

The simple equations that govern the operation of the sensor are:

$$\begin{aligned} R_{sensor} &= R_{cold} + \tilde{R} \\ &= R_{cold} + \alpha(T - T_o) \end{aligned} \tag{4.1}$$

$$\begin{aligned} P_{sensor} &= I_{sensor} R_{sensor}^2 \\ &= I_{sensor} (R_{cold} + \tilde{R})^2 \end{aligned} \tag{4.2}$$

where $\tilde{R}$ is the change in resistance that occurs with current flow and is related to the change in temperature of the sensor by the parameter α which also known as the temperature coefficient of resistance (TCR). Note that Equation 4.1 assumes a linear dependence between the temperature and the resistance. In practice, however, this is not true and generally higher order terms, e.g., $\beta(T - T_o)^2$ and $\gamma(T - T_o)^3$, are needed for adequate modeling of the behavior.

The physical details of the sensor are provided to give the reader a basic understanding of the device construction. The 150 μm long, 3 μm wide, 4000 øA high

[1]This basic principle has been in use for over 100 years in the field of hot-wire anemometry. The only difference being that most standard hot-wire anemometry relies on small hand-crafted platinum wires for the resistive element, while we rely on microfabricated polysilicon wires.

polysilicon wire sits on a 200 μm square, 1.2 μm thick silicon nitride diaphragm. This diaphragm covers a 2.0 μm deep vacuum cavity which serves to minimize thermal losses to the silicon substrate. The thermal losses to the substrate and the metal contacts at either end of the polysilicon resistor ultimately limit the sensitivity of the sensor. These additional paths for heat transfer also give the sensor multiple thermal time constants which limit the frequency response[2]. A detailed picture of the sensor (Figure 4.2) shows the various parts and the relative geometry. The exact sensor width, length, and height are chosen empirically to optimize sensitivity, noise, matching constraints and amount of flow disturbance [30].

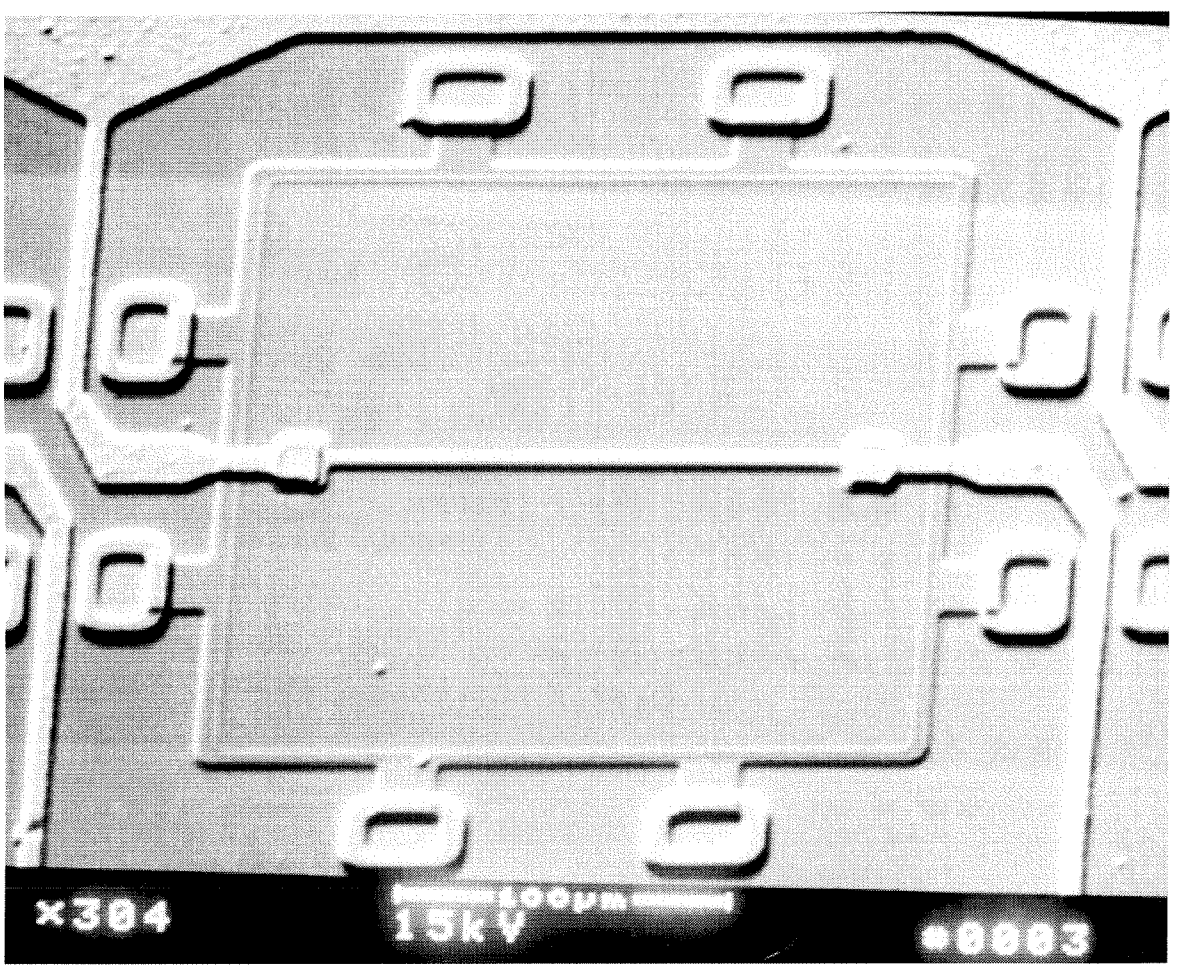

Figure 4.2: SEM photograph of the μ-shear stress sensor which shows the polysilicon heater element, Si_3N_4 diaphragm and polysilicon to metal contacts for circuit connectivity.

[2]While also making the circuit analysis much more involved.

4.2.1 Thermal time constant

The sensor can be thought of as having a thermal time constant characterizing the rate at which heat is transferred from the element to the fluid. Often the sensor has multiple thermal time constants because of the different heat transfer paths from the element, e.g., the Si_3N_4 diaphragm, the metal contact, and the silicon substrate. However, in practice we only consider the transfer of heat from the polysilicon wire to the flow as the dominant thermal time constant. This thermal time constant is referred to as M in the equations. For the sensors that we use, M is approximately 0.5 to 0.7 ms.

4.2.2 Electrical Properties

From an electrical point of view, the sensor is a simple two-terminal resistor. If one end of the sensor is tied to ground, a n-sensor array can be built using only $n + 1$ connections. An array of 25 such sensors, each having one terminal connected to a common ground, is depicted in Figure 4.3.

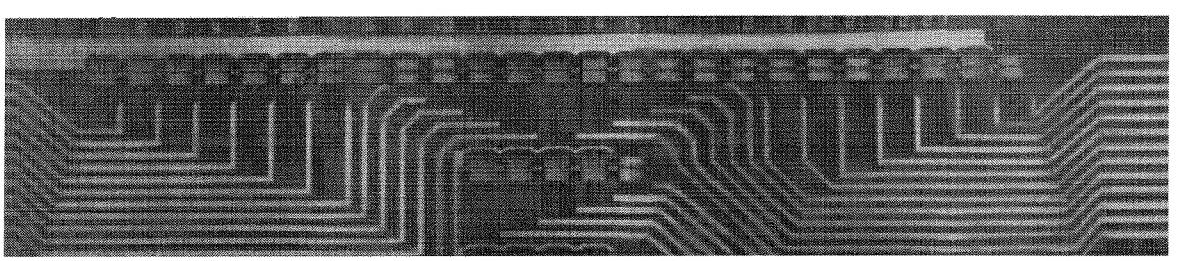

Figure 4.3: Micro-photograph of an array of 25 μ-shear stress sensors. The individual sensors are spaced 300 μm apart for a total array width of 7.5 mm.

4.2.3 Discrete and Integrated Sensors

Currently we use two types of sensors categorized as Type I and Type II. Both types of sensors have similar characteristics, the difference between the two being that Type II sensors are intended for fabrication in a wafer-scale integration setting while Type I are intended for fabrication on a separate substrate for use in discrete

applications. The Type I sensor has a lower resistance, and hence which means a different current versus voltage characteristics[3]. Most of the data in this thesis was collected using discrete arrays of Type I sensors. The discrete sensor characteristics are reviewed in the next subsection.

4.2.4 Electrical Properties of the Type I Sensor

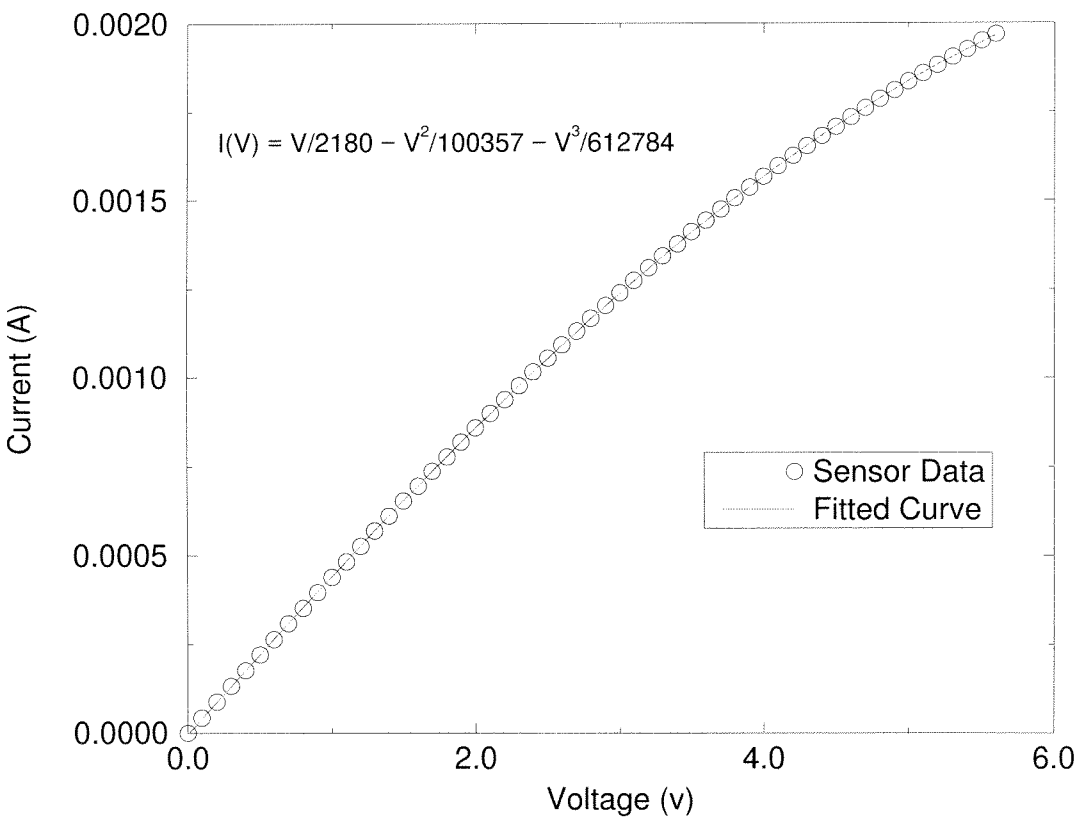

Figure 4.4: Current versus voltage plot for the Type I μ-shear stress sensor with no flow applied. The sensor data is fit with the third order polynomial shown.

The cold resistance (R_{cold}) of the sensor is about $2300\Omega \pm 10\%$. As more voltage is applied across the sensor, the *effective* (operating) resistance increases. This means that the sensor has a positive temperature coefficient of resistance. The current

[3]For more details, please see Section 7.3.1.

versus voltage curve for this sensor is given in Figure 4.4. We note that the slope of the curve decreases with larger applied voltage meaning the resistance is increasing with with increasing power input. The curve can be fit accurately using a third order polynomial.

$$I(V)_{fitted} = \frac{V}{2180} - \frac{V^2}{100357} - \frac{V^3}{612784} \tag{4.3}$$

The primary reason for obtaining the fit is to provide an equation within a SPICE[4] model for reasonable electrical simulations of the sensor. In Equation 4.3, we note that the denominator of the first term of the polynomial fit is equal to the cold resistance R_{cold} of the sensor. This is because the linear term fits the curve best at very low voltages. This means that small currents and, hence, lower power is dissipated as heat[5].

There are two ways to utilize the current versus voltage curve for sensing. The first way involves applying a known fixed potential (or current) to yield a unique current (or voltage) giving a nominal resistance. As the sensor operates, the changing resistance creates a new value of current (or voltage) which can be sensed. The second mode of operation is to apply feedback to the circuit that drives the sensor resistor. This method permits selection of an operating resistance causing the feedback to find the appropriate voltage and current which satisfies the circuit constraints. Both of these methods of driving the sensor are covered in the next two sections. We note that care must be taken not to apply excessive voltage (or pass too large a current) to prevent the sensor from overheating and burning out.

4.3 Constant Current Operation

One of the simplest ways to operate the μ-shear stress sensor is to pass a fixed current through the sensor and monitor the voltage developed across the sensor. This is called a constant current sensor driver stage. This circuit can easily be im-

[4] An electrical simulation program.

[5] This implies a lower operating temperature close to T_0, i.e., the cold temperature.

plemented with a CMOS current source. Similar CMOS current sources are used in operational amplifiers and other circuits to supply fixed currents into a variable load. The simplest circuit that acts as a current source is a current mirror as depicted in Figure 4.5. This limitation of simple current mirror approach is that it exhibits poor current regulation with changing output voltage. Namely, the Early voltage[6] of the devices can be low, i.e., $V_{an} \approx 30\text{V}$ and $V_{ap} \approx 20\text{V}$ for typical sized devices. This contributes an error to the current term that can be seen in Equation 4.9.

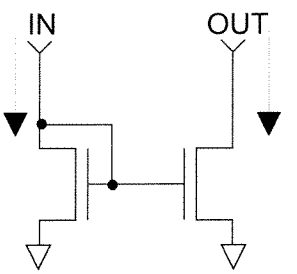

Figure 4.5: Schematic of a CMOS current mirror which can be used as a current source to drive the μ-shear stress sensor.

We have fabricated a number of different current sources in order to achieve a useful driver stage for the μ-shear stress sensor. The topologies of these current sources are similar in an attempt to achieve a constant current over the widest range of input voltages when working from a 5V power supply. The operation of the sensor is as follows:

$$V_{out} = I_{const} R_{sensor} \tag{4.4}$$

$$= I_{const} R_{cold} + I_{const} \tilde{R} \tag{4.5}$$

If only the differential terms of this equation (i.e., the non-varying terms are ignored) are considered, the measured value is

[6]The Early voltage is the x asymptote of the current versus voltage curve of a MOSFET device in saturation. It is related to the current $I_{\text{ds}}$by the formula $I_{\text{ds}} = I_{\text{ds}}(V_d = 0)\frac{V_{\text{ds}}}{V_{\text{a}}}$. The parameter λ is defined as $\frac{1}{V_{\text{a}}}$.

$$\tilde{V}_{out} = I_{const}\tilde{R} \tag{4.6}$$

This equation produces an output that is exactly proportional to the change in resistance (the signal of interest).

The error which results if I_{const} is a function of V_{out} can be estimated assuming a linear error relationship (ζ). This accounts for the non-idealities in the current source.

$$I_{const}(V_{out}) = I_{const} + \zeta V_{out} \tag{4.7}$$

$$\tilde{V}_{out} = I_{const}\tilde{R} + \zeta V_{out}\tilde{R} \tag{4.8}$$

$$\frac{\partial \tilde{V}_{out}}{\partial \tilde{R}} = I_{const} + \zeta V_{out} \tag{4.9}$$

The error is equal to ζV_{out}. The factor ζ for a simple current mirror is equal to λ.

For the different current sources that we have built, the Early voltage (V_a) of a MOS output device sets the upper bound for the dependence of current on the output voltage of the current source. With the use of series feedback, we have increased the effective Early voltage by a factor equal to the feedback gain. For a current source that incorporates feedback (Figure 4.6), ζ is reduced to $\frac{\lambda}{A_o}$ where A_o is the DC feedback gain. The reduction in ζ leads to an improvement in the performance of the constant current sensor driver. A typical value of λ for a n-channel device is 0.03. A typical value for A_o for a simple transconductance amplifier is about 100 which leads to a typical $\zeta = 0.0003$.

4.4 Constant Temperature Operation

The constant temperature mode of sensor operation keeps the resistance of the sensor constant. Because we assume the relationship (to first order) between temperature and resistance to be given by Equation 4.1, if the resistance can be kept constant under changing operating conditions the temperature at the sensor should

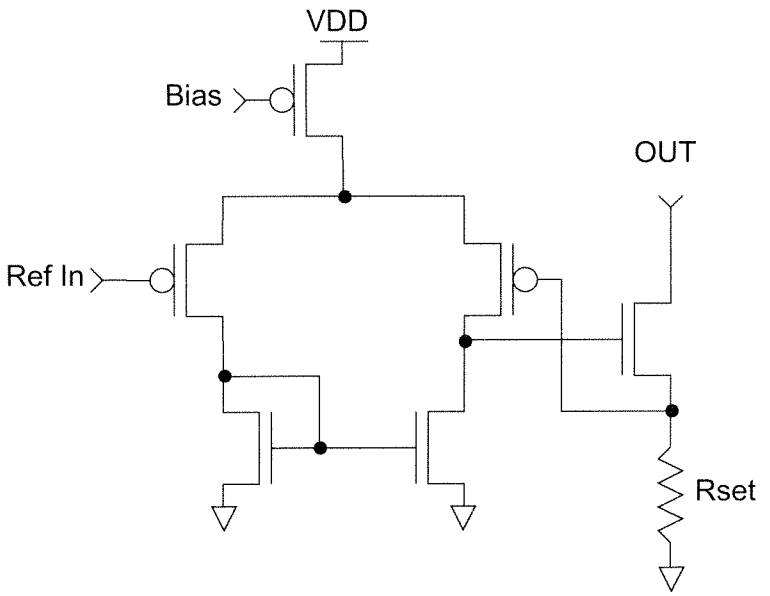

Figure 4.6: Schematic of a current source used to drive the μ-shear stress sensor which boasts an improved current regulation. The current source has a value equal to $\frac{\text{Rset}}{\text{V(Ref In)}}$. Rset is constructed from polysilicon and has a value of about 175 Ω.

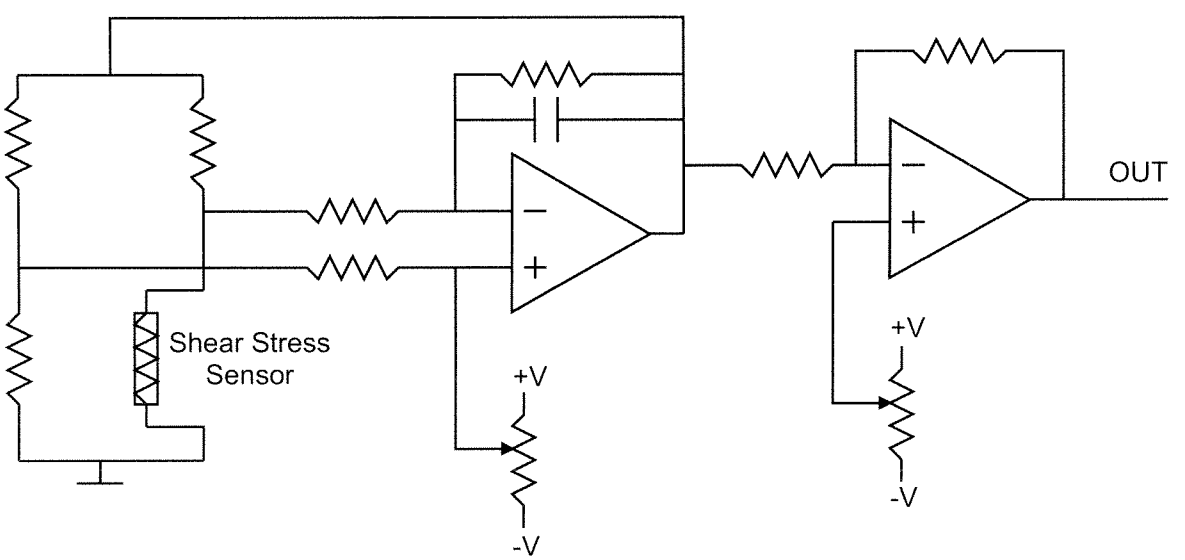

Figure 4.7: Schematic of a constant temperature circuit constructed using discrete components. The second amplifier merely serves to amplify the output of the first amplifier by a factor of 10.

correspondingly remain constant. In the case of our μ-sensors, the normal operating temperature for the Type I sensors reaches about 125°C and is similar for Type II sensors.

The constant temperature (CT) circuit places the sensor in one leg of a Wheatstone bridge configuration (Figure 4.7). The operational amplifier (opamp) output terminal is connected to the top of the bridge which converts the output voltage to a current. The high open-loop gain of the opamp forces the imbalance between the two sides of the bridge to be as small as possible through the use of feedback at the top of the bridge.

Because one side of the bridge contains resistors in a fixed ratio, the feedback strives to keep the variable sensor side regulated to the same ratio as the constant side. The top resistor in the side of the bridge containing the sensor sets the operating resistance of the sensor if the two resistors in the other side of the bridge are equal. The overheat ratio (a_R) is the ratio between the desired operating resistance and the cold resistance (R_{cold}) of the sensor. We typically run all of our constant temperature circuits at an overheat ratio of 0.1, which means that $R_{operating} = 1.1R_{cold}$. The opamp output voltage is an error signal that drives this bridge as well as serves as the output signal of the circuit. The output varies inversely to the change in incremental temperature.

More details of the constant temperature circuit are presented in Appendix A.

Sensor Voltage To Shear Stress Calibration

The output of the CT circuit is related to the surface shear stress by the following equation derived from King's Law,

$$E^2 = A + B\left(\tau_w\right)^{\frac{1}{3}} \tag{4.10}$$

where E is the output voltage, and A and B are constants that are determined from a fitting of a calibration curve. This calibration involves using an additional equation for a channel flow setting [26]:

$$\frac{U_\tau}{U_c} = 0.1079(Re)^{-0.089} \tag{4.11}$$

Integrated Constant Temperature Driver

An integrated version of the constant temperature circuit has the following advantages over the discrete version:

- lower capacitance on the circuit nodes, especially the sensor which permits faster response
- lower power requirements on the opamp (+5v V_{dd}versus $\pm$ 15v V_{dd}) which saves system power
- shorter wire lengths which improve noise immunity in the environment

More results and further discussion of the integrated constant temperature circuit appear in Section 7.3.1.

4.5 Yaw Response

The sensor responds best to flows perpendicular to the sensor orientation. At this orientation the heated cross-sectional area of the sensor has maximum breadth to the flow. This corresponds to a yaw angle of 0^o where the yaw angle is measured between the flow and the perpendicular axis of the sensor in the surface plane. Rotation of the sensor is expected to follow a cosine law [44]. The plot of yaw angle versus response is given in Figure 4.8. The important thing to note is that the response decreases significantly as the sensor orientation is changed. This orientation sensitivity means that the sensor is optimized for use as a single flow-line sensor. Although there are sensor topologies called x-wires that address this shortcoming, none of the designs have yet been used in either a discrete array or an integrated version.

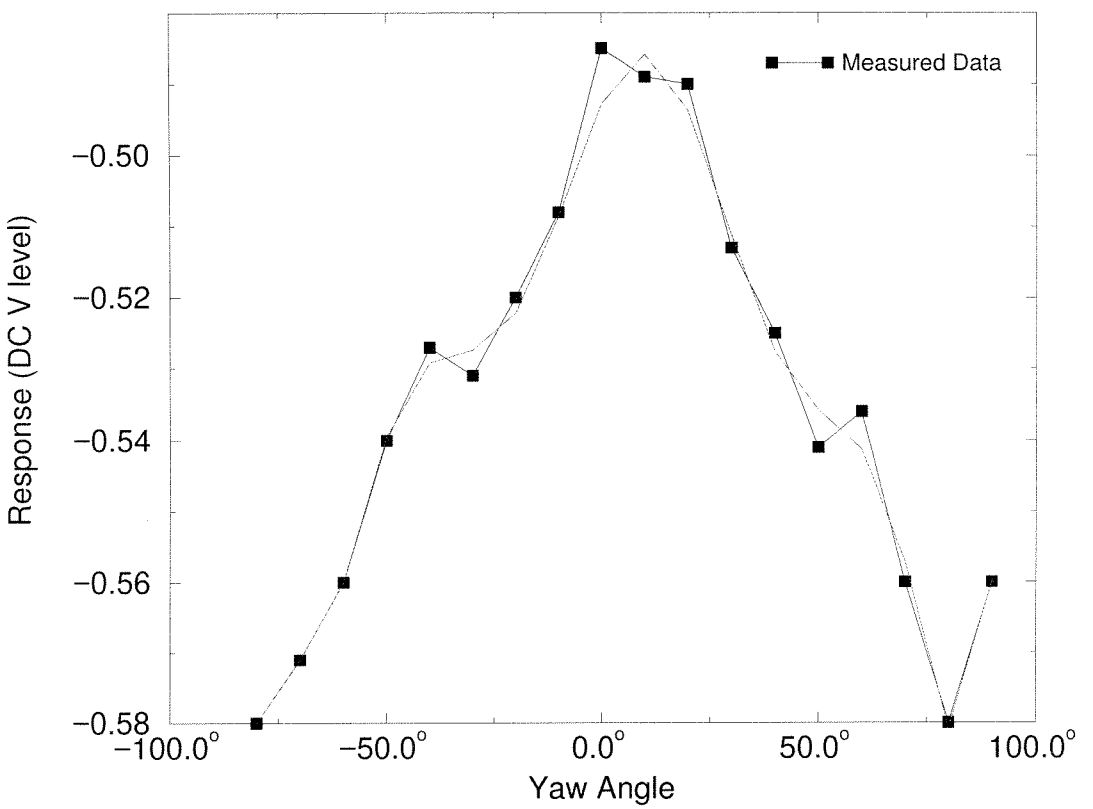

Figure 4.8: Yaw response of the μ-shear stress sensor plotted for various orientations in the wind tunnel with a centerline velocity (U_c) of 10 $\frac{\mathrm{m}}{\mathrm{s}}$.

4.6 Microactuator

The basic idea behind the μ-actuator is to build a flap that can impart a force normal to the surface. Since we want to actuate regions on a similar scale to the vortices, it is only natural to construct actuators using micromachining principles. The μ-actuator is essentially a thin plate of semiconductor material that can be actuated via either a thermal or magnetic method. The actuators that have been constructed for this project have varied greatly in their design. The actuator is generally constructed with at least one metallic coil which allows current flow to create a small magnetic field. Depending on the resistance of the coils and return path, the current flow can also create a significant amounts of heat. The resulting thermal stress can also cause the actuator to bend. The actuator can achieve up to a 200 μm vertical deflection. Figure 4.9 shows some of the μ-actuators.

However, the current μ-actuators cannot attain sufficient mechanical bandwidth necessary for our application. In theory, the resonant frequency of the actuator plate should be on the order of 3 kHz. However, at the present time only a 100 Hz bandwidth has been physically achieved. Bulk micromachined actuators are being investigated as a alternative way to provide the necessary bandwidth and deflection for use in the turbulent boundary layer. Finally, the surface micromachined flaps do not seem to have the structural integrity necessary to stand up to the forces encountered in turbulence, and break for flows in excess of 10 $\frac{\mathrm{m}}{\mathrm{s}}$.

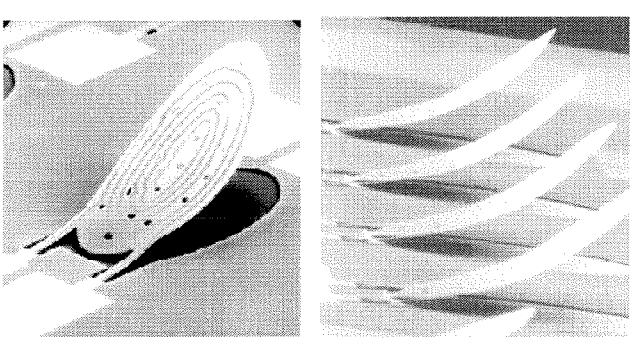

Figure 4.9: μ-actuator array is pictured. This particular array is a surface micromachined version. The inherent stress in the plate causes the flap to initially bend up.

Chapter 5 VLSI

Analog circuitry gives us the ability to build microelectronic circuits which perform the real-time processing necessary in building our integrated system. A natural question to consider is whether or not analog-domain processing is the best method to process our signals.

5.1 Processing Methodology

Signal processing can be performed using several different methods. The different methods correspond to the circuits used to do the processing. The two basic types of processing are analog and digital. These two types of processing can be further delineated based on how often the data is processed. A list of the different types of potential analog and digital signal processing methods we consider are as follows:

1. continuous time analog processing
2. sampled data analog processing
3. multiplexed and sampled data analog processing
4. input-dedicated digital processing
5. multiplexed digital processing

Clearly, the two types of digital processing listed are normally considered sampled data systems. Under certain area, power, or circuit complexity constraints, multiplexing of signals must be considered. The key to the best approach is to consider the statistics of the input signal, the requirements for processing, the noise that can be tolerated and finally system constraints such as die area and power consumption.

Very simply stated, the surface shear stress inputs to our system are the product of a very complex and unpredictable real world process. Consequently, these signals are, by consequence, continuous-time and nearly-stochastic. As a result, the processing needs to be either continuous-time or at a sufficiently high data sampling rate to support discrete time processing along with the possibility of multiplexing the signals.

5.1.1 Input Bandwidth

In our system, we consider our usable sensor input bandwidth to be approximately 10 kHz (see Figure 1.8). If we wanted to do a reasonable job of processing this data in the sampled data domain, the sampling rate must be, at a minimum, twice the input bandwidth[1]. In practice, this lower bound on the input bandwidth is often insufficient. Oversampling factors of five or ten times are used to retain some additional information. Furthermore, if significant signals exist at higher frequencies the problem of aliasing is introduced[2]. To combat this problem, either expensive multiple pole filters are used or an even higher rate of sampling followed by a simpler filter is employed. Having more bandwidth at the start may improve the signal to noise ratio, similar to an oversampled delta-sigma analog to digital (A/D) converter, which is beneficial if further steps of lossy processing are performed. The conclusion this reasoning is that a sampling circuit capable of resolving the dynamic range necessary for the required sampling rate at each sensor output.

5.1.2 Digital Processing

Approaching this problem in the digital domain without multiplexing would require an analog-to-digital converter at each sensor. The purpose behind multiplexing is to share an expensive resource where possible. In our case, one expensive

[1]This constraint is derived from the Nyquist's theorem. The theorem states that the rate of sampling at which reconstruction of the input signal preserves necessary information should be at least twice the highest frequency of concern.

[2]Aliasing is where high frequency information is "folded" back into lower frequencies by the process of sampling.

resource would be the A/D converter. However, the tradeoff then becomes building fewer A/D converters that run at a higher frequency. In the case of our 20kHz sampling requirement, choosing to multiplex 10 sensor signals onto one converter raises the converter's sampling frequency from 20kHz to 200kHz. While building a 200kHz, n-bit converter is not an issue for this thesis, other considerations are important. The first concern is the precision (n-bit) of the converter. If n is equal to 8 bits, the converter complexity (in both area and power) is reasonably manageable. However, if n is 12 bits then the converter design becomes substantially more difficult[3]. Another concern when using a high speed A/D converter whose input is multiplexed over many sensors is one of signal routing both from the sensor to the converter and from the converter to the next stage of processing. Analysis of such a routing scheme shows that the topological nature of the processing is not handled well by a centralized processing scheme (Figure 5.1). Because we would like to extend the circuitry to handle a wafer-scale number of sensors and actuators, the necessary wiring and communication pose a serious concern.

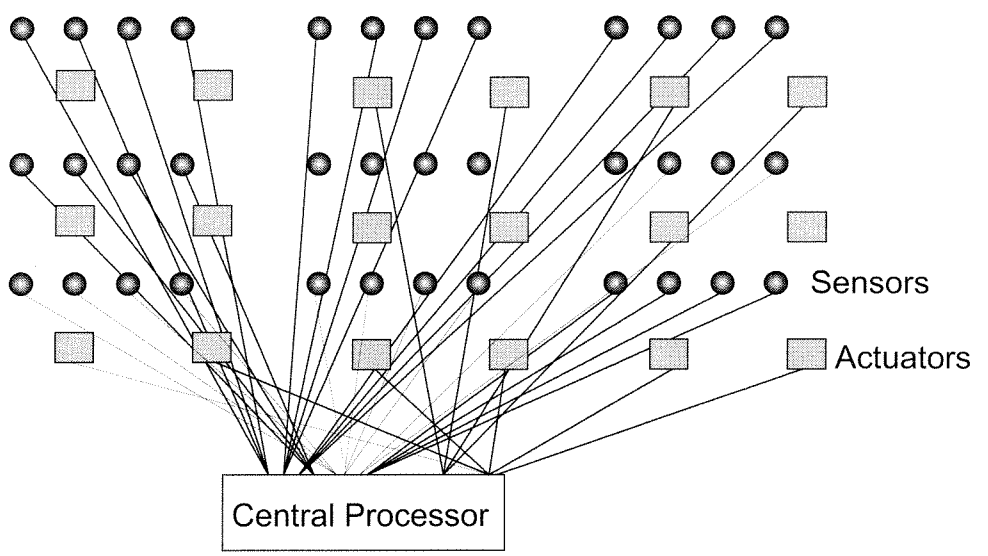

Figure 5.1: Diagram of the connectivity required to enable a central processing unit to handle a two dimensional array of sensors and actuators.

[3]It is not uncommon to see 8-bit flash A/D converters. However, 12-bit flash converters are virtually unheard of due to the physical matching constraints in generic CMOS precluding the use of 4096 well matched components.

Sampled-time analog systems, with or without multiplexing, have similar complexity to digital systems. Even though an A/D converter is unnecessary, a good analog storage methodology now becomes a constraint. In fact, the storage requirements are daunting, especially if the algorithm performs both spatial as well as temporal processing. A good exemplification of this problem is the filtering of an audio-band signal[4].

5.1.3 Distributed Real-time Analog Processing

The distributed continuous-time analog approach is simple. First, since the real world signals are continuous time, they are most naturally handled by continuous time circuits. Second, the topological nature of the problem leads one to build a distributed system. Since the regions of high shear stress occur randomly in space, the system need to be able to process the information anisotropically. Third, because of the large relative size of the sensors and especially the actuators to the basic circuitry, the chip area necessary to build large central processing units or digital circuitry is unavailable[5].

5.2 Analog Processing Methodology

Given the elongated aspect ratio and temporal unpredictability of the streaks that cause high shear stress, we can build a specialized circuit to detect these regions. A circuit that meets our criteria for detecting the high shear stress regions and then outputting an appropriate control response can be designed using familiar circuit blocks. We have built our system to detect the regions of high shear stress by applying simple principles of engineering to the signal analysis problem.

[4]For a simple two pole audio band filter, the circuit consists of a number of capacitors and opamps. The total area for this circuit would approach 1 mm^2 for all the active and passive components.

[5]Also the constraints placed on our system by the relative size and placement of the actuators limits the number of sensors and size of the associated processing circuitry.

5.2.1 Neural Network Justification

The CFD simulations (Section 3.1.1) provide an idea of how to predict the wall actuations necessary to minimize surface shear stress. In the CFD simulation experiments, an effective pattern detector was the span-wise derivative of the surface shear stress (Section 3.1.3). This result indicates that the pattern detector responded with a large control signal whenever the values from neighboring grid points differed substantially. Hence a reasonably simple initial approach is to actuate whenever a large span-wise derivative is found.

5.2.2 Intermittency Function Justification

An additional justification can be found for our approach in a similar signal processing scheme presented by Bruun (Figure 5.2) [9]. His depiction of the turbulence detection system shows the different stages in which the information is processed temporally to detect turbulent behavior. In addition to this basic system, ours is modified by using the spatial turbulent signal detector to trigger an output response.

5.3 Information Flow

Figure 5.3 depicts a detailed schematic of a single column (or channel). The CT sensor output signal feeds into a pre-amplifier stage (Figure 5.4). A buffer (Figure 5.6) distributes the differentiated and amplified signal to a non-linear resistive network composed of horizontal resistors (HRES) circuits [42] (Figure 5.8). The spatial filtering constants between the columns are chosen to reinforce activity within a column and discourage activity between adjacent columns. The filtered signals feed to a symmetric anti-bump circuit [15] (Figure 5.9) whose operation mimics a soft comparator with a somewhat adjustable dead zone. The function of the circuit is to indicate when a particular column has registered a large shear stress value, while the neighboring columns have not. This is our topological constraint which func-

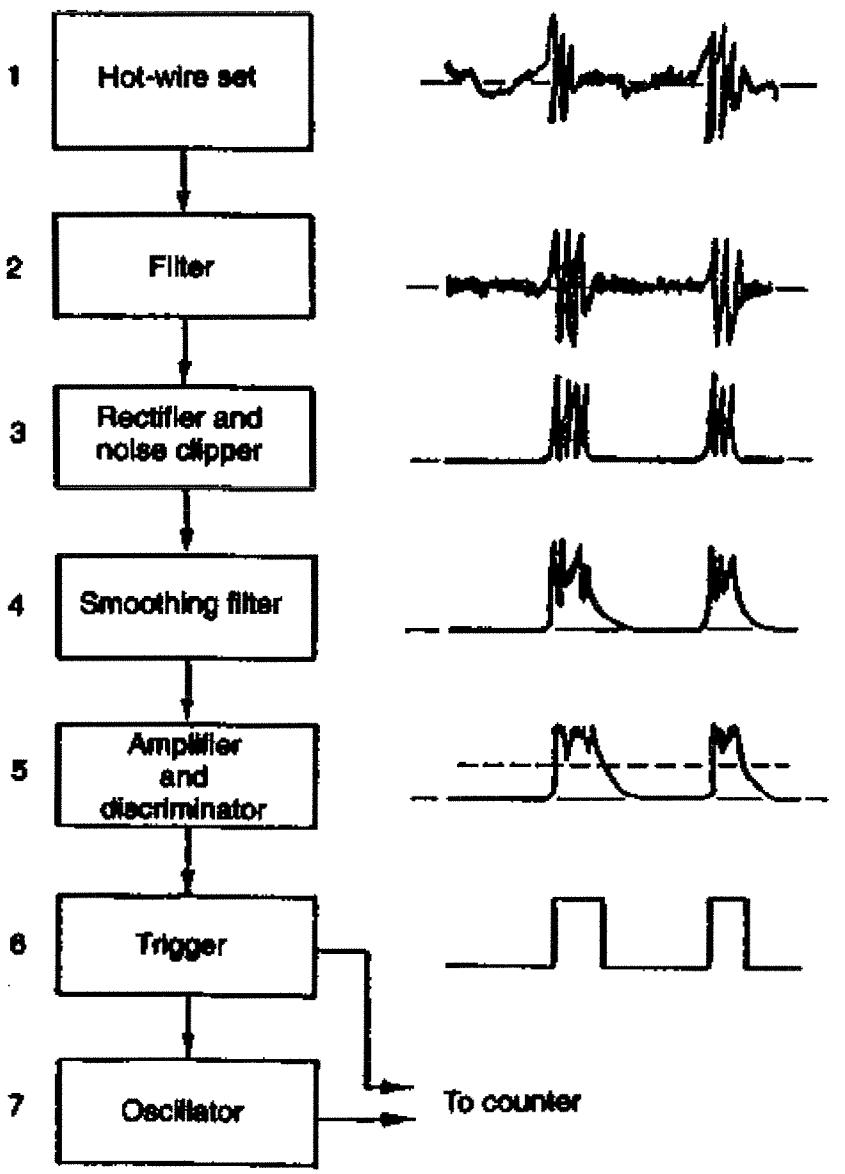

Figure 5.2: Figure of the signal processing system described by Bruun to detect turbulence using a hot-wire signal as the input. The figure is reproduced from [9].

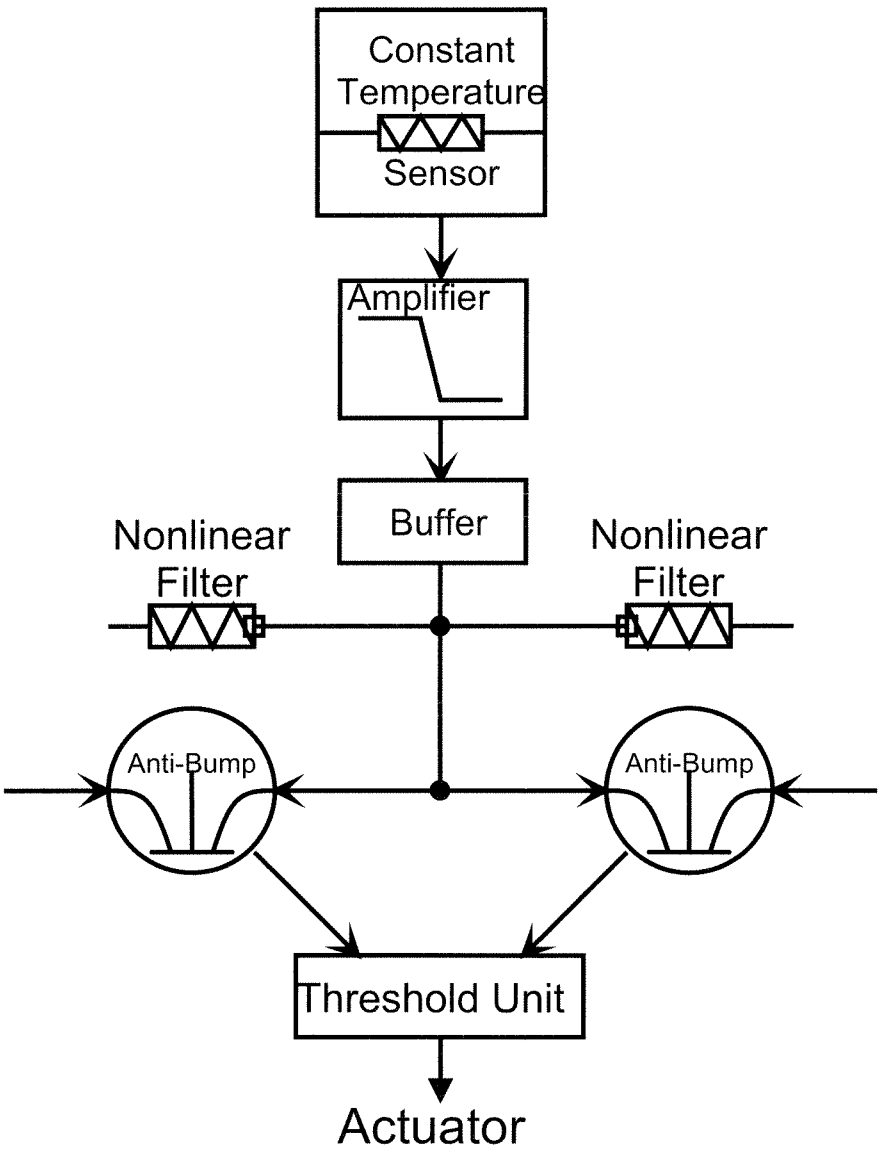

Figure 5.3: Diagram showing one channel of the detection/control information processing flow.

tions a approximate spanwise differentiator. The output current of the anti-bump circuit is compared to a threshold using the actuator drive circuit in Figure 5.10. If the value exceeds the threshold, the circuit triggers the actuator by turning on a pull-down transistor which allows current to flow through the actuator.

5.4 Preamplifier

The preamplifier has several different roles which we shall examine in greater detail. First, it removes any DC offset in the CT sensor signal output. This is useful because the sensor resistance matching is only within 10% at best, hence the DC output of CT driver will also vary similarly. The variation in sensor resistance can saturate circuits that have constant passive components and high gain.

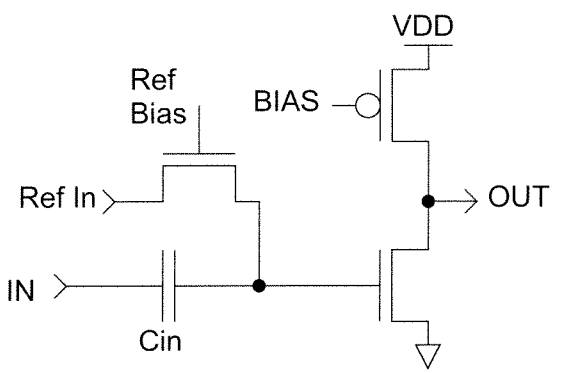

Figure 5.4: Schematic of the pre-amplifier used in our system. The circuit accomplishes several different functions described in the text (Section 5.4).

By applying the CT sensor signal into a capacitor, the DC component can effectively be removed. However, a reference DC signal should replace the removed DC component in order to correctly drive the single-ended amplifier that follows in the signal path. The inherent nature of a non-differential input amplifier, operated either open-loop or with feedback, requires a reference voltage in order to provide a reasonable output signal for small input amplitudes.

Applying the reference voltage (V_{ref}) accomplishes two things with this one circuit. First, we restore a consistent DC level to all the preamplifiers. Second, we adjust the corner frequency of the zero in order to get the desired filtering. The corner

frequency of the zero is determined by the impedance across which we apply the V_{ref} voltage and the input capacitor. Ideally we would like this zero to be at as low a frequency as possible in order to remove the DC component while allowing signals of interest at higher frequencies to pass unattenuated or with only a constant phase shift. We have used two different circuits to create this impedance. The first is simply a transistor that is turned on very weakly. In the deep subthreshold region the drain to source impedance can be as high as 10 MΩ. The other impedance method employs a "Tobi"-element which consists of a PMOS device connected to have a MOS diode characteristic in one direction and a diode characteristic in the reverse direction [15]. The desired corner frequency is between 1 and 100 Hz. [6]

It is important to mention that this functionality of the preamp is similar to that of a signal differentiator (in time). The differentiator has a response:

$$\frac{V_o}{V_i} = (1 - \frac{s}{z_1}) \tag{5.1}$$

The gain stage that follows the capacitor and impedance is a single transistor common-source amplifier with passive current source load. This amplifier is a pure class A amplifier. The whole circuit has a simple frequency response ($PreAmp$):

$$\frac{V_o}{V_i} = \frac{A_o\,(1 - \frac{s}{z_1})}{1 - \frac{s}{p_1}} \tag{5.2}$$

where $A_o = g_m(r_{o_n} || \, r_{o_p})$ is the DC gain (approximately 200 for the circuit) and p_1 is the dominant pole of the amplifier. This amplifier serves to increase the amplitude of the CT shear stress voltage signal. Even though the amplifier is non-linear (Figure 6.1), it performs the task for the small signals of interest. With the amplifier bias in the weak inversion region[7], we get sufficient amplification with minimum current usage through the class A amplifier. To further increase the gain we can

[6] In practice, this corner frequency varies between $\approx$ 1 and 100Hz for the two different kinds of impedances that we have used to build our circuits.

[7] Which is also called the subthreshold region of operation where the primary current flow is due to diffusion, not drift.

add a cascode transistor (Figure 5.5) giving $A_o = g_m(r_{o_n1}(1 + g_m r_{o_n2})||r_{o_p})$ (about 1000).

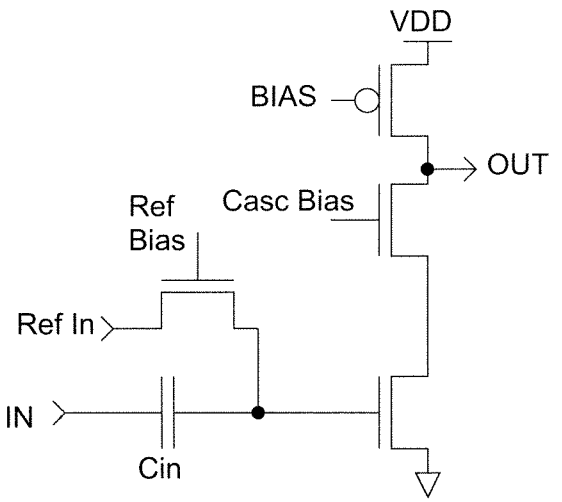

Figure 5.5: Schematic of the cascode-enhanced pre-amplifier in which the cascoding enhances both the gain and bandwidth while reducing the available output voltage swing.

5.5 Buffer

Once the CT shear stress signal has been amplified, we need to buffer the resulting signal since our preamplifier is not designed to drive a significant resistive or capacitive load. The buffer design we use provides both rail to rail input and output capability [34]. The gain of this amplifier is given by $(Buffer)$:

$$R_o = r_{o_n}||r_{o_p} \tag{5.3}$$

$$G_m = G_{m_n} + G_{m_p} \tag{5.4}$$

$$\begin{aligned} A_v &= G_m R_o \\ &= (G_{m_n} + G_{m_p})(r_{o_n}||r_{o_p}) \end{aligned} \tag{5.5}$$

As an aside, we have modified the amplifier by adding transistors to shunt the input pair bias current to give us a flat transconductance between the input and

output, see Figure 5.7. This amplifier has $G_m = \text{const}$ and, therefore, is not as dependent on input common mode voltage as the normal buffer.

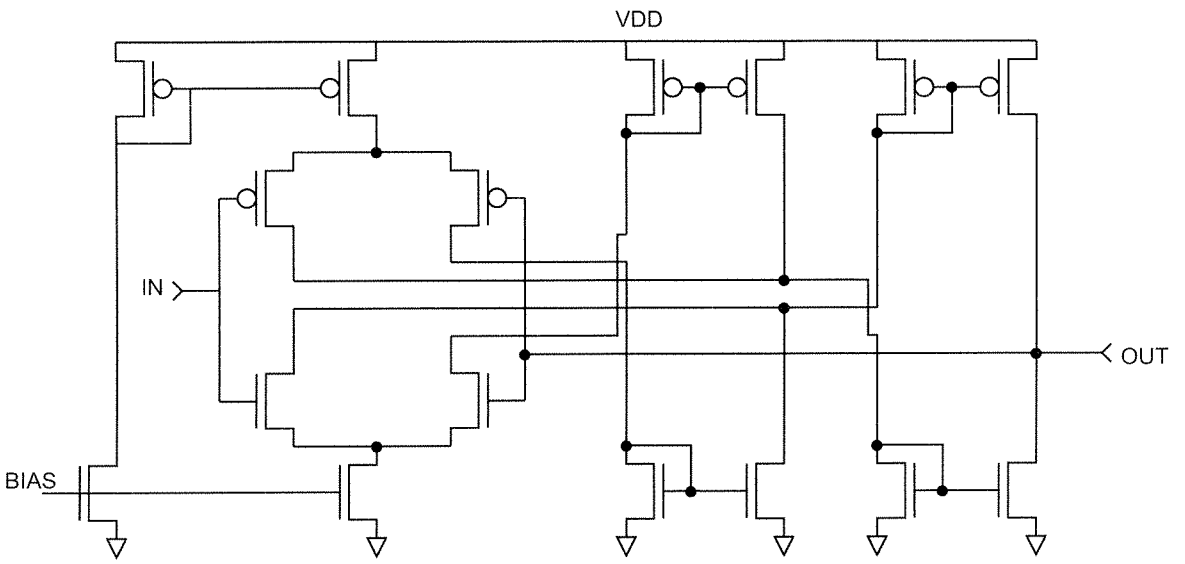

Figure 5.6: Schematic of the buffer amplifier from [34]. The amplifier is very useful because it has both rail-to-rail input and output common-mode ranges.

5.6 Horizontal Resistor (HRES)

Once the signal is buffered, we need to smooth out the signal to a certain extent. The rationale behind this is that the noise floor has been raised in the sensor signals due to amplification. The filtering helps is avoiding falsely detecting a non-existent difference between shear stress signals caused by the noise. However, the filter must be design so as not to remove large differences in the input as they surely represent salient signals. The HRES (or, for that matter, any other spatial filtering circuit) is a good candidate for this function. An important issue with the HRES is the small tunable range over which the current versus voltage characteristic is unsaturated, however in this application this is a desirable feature. More specifically, when topologically close values are divergent we allow the HRES to "break," that is we do not smooth away these differences on which we want to act. The $Filter$ equation is:

$$I_{AB} = \frac{I_B}{2} \tanh \frac{V_A - V_B}{2U_T} \tag{5.6}$$

$$\begin{aligned} G_{effective} &= \frac{I_{AB}}{V_A - V_B} \\ &= \frac{\frac{I_B}{2} \tanh \frac{V_A - V_B}{2U_T}}{V_A - V_B} \end{aligned} \tag{5.7}$$

where I_{AB} is the current between the resistive terminals at potentials V_A and V_B, and I_B is the bias current in the HRES.

5.7 Anti-Bump Circuit

The next stage of processing imposes our topological constraints. We use the anti-bump circuit (Figure 5.9) to give us a gentle comparison. The anti-bump circuit is one of the class of circuits built around a signal correlator [15]. The output signal is proportional to the difference between two inputs according to a function which non-linearly rewards only dissimilarity. If the two input values are not widely divergent, then the current output is negligible.

This property is the basis for the high shear stress region detection. The range over which the anti-bump circuit outputs a current is somewhat adjustable. By using a completely symmetric circuit, we can extend the working region of the circuit from the order of several $U_T = \frac{kT}{q}$ up to the order of a volt or two. The $AntiBump$ function is:

$$I_R = \frac{I_B \tanh \frac{V_R - V_L}{2U_T} + I_B}{2 + \frac{S}{2} sech(\frac{V_R - V_L}{2U_T})} \tag{5.8}$$

where I_R is the current output of the right side given a voltage difference V_R and V_L between the right and left side inputs and S is an effective sizing ratio of the transistors.

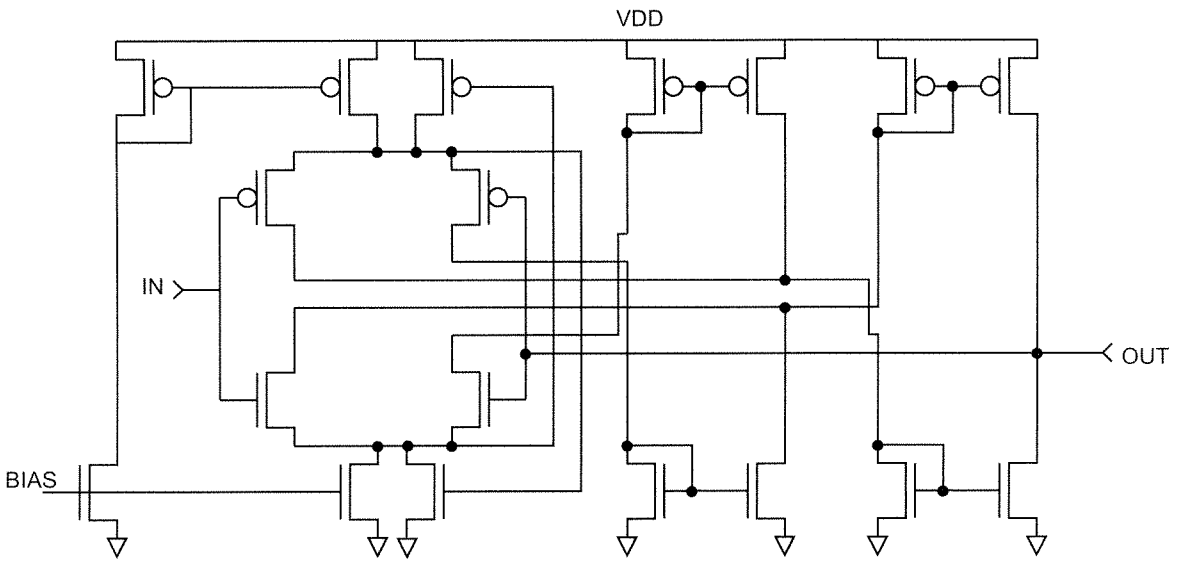

Figure 5.7: Schematic of the modified buffer amplifier that maintains a more steady G_m across a large common-mode input range.

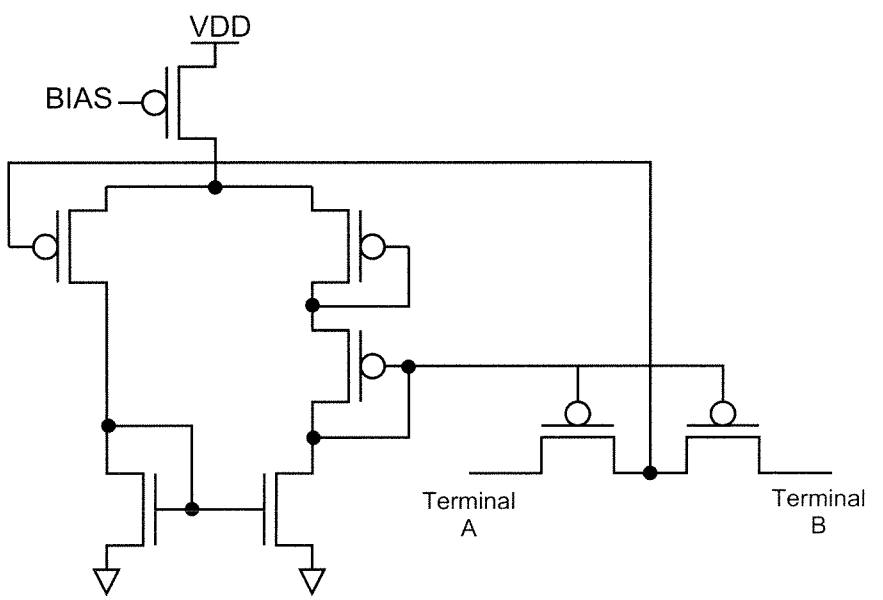

Figure 5.8: Schematic of the HRES circuit showing the transistors that act as the non-linear resistive elements along with the bias circuitry that controls the resistive transistors' gate potential.

5.8 Actuator Driver

The purpose of the actuator driver is to compare the current output of the anti-bump circuit representing the value from one column to a threshold and to activate the μ-actuator associated with that column. The activation of the μ-actuator is accomplished by turning on a transistor with a large $\frac{W}{l}$ ratio to act as a current sink. Because the actuators require very large currents to get a reasonable actuation, the large transistor is required to provide a very low impedance compared with the actuator. The $PostAmp$ function is the following:

$$I_{OUT} = g_{m_N} g_{m_P} (r_{o_N} || r_{o_P}) \frac{U_T}{\kappa} ln(\frac{I_{IN}}{I_O}) \tag{5.9}$$

The present method of actuation is to simply turn on the actuator when the control signal is active. Considering the current maximum bandwidth limitations of the actuator, we are unable to verify more complicated driving schemes at the present time. From a fluid mechanicist's point of view where the ideal blowing or suction is unachievable to the precision obtainable in the CFD simulation, it is difficult to apply many fluid mechanics principles to aid the selection of an actuation algorithm. However, there are a few principles that can provide some guidance. First, introducing mass at the point of actuation is undesirable because it would be difficult to construct such a system. Second, creating more μ-vortices by influencing a relatively calm flow will also have a deleterious effect because it would create more surface shear stress. The actual optimum actuator driving signal is further complicated because the actuator design is also in a state of flux (Section 7.3.2).

5.9 Complete System

The complete system that we have built[8] is schematically depicted in Figure 5.3. The flow of information through the circuit is governed by the equations of the individual circuits presented in the previous sections. The composite circuit equa-

[8]For historical reasons it is called the control/detection chip.

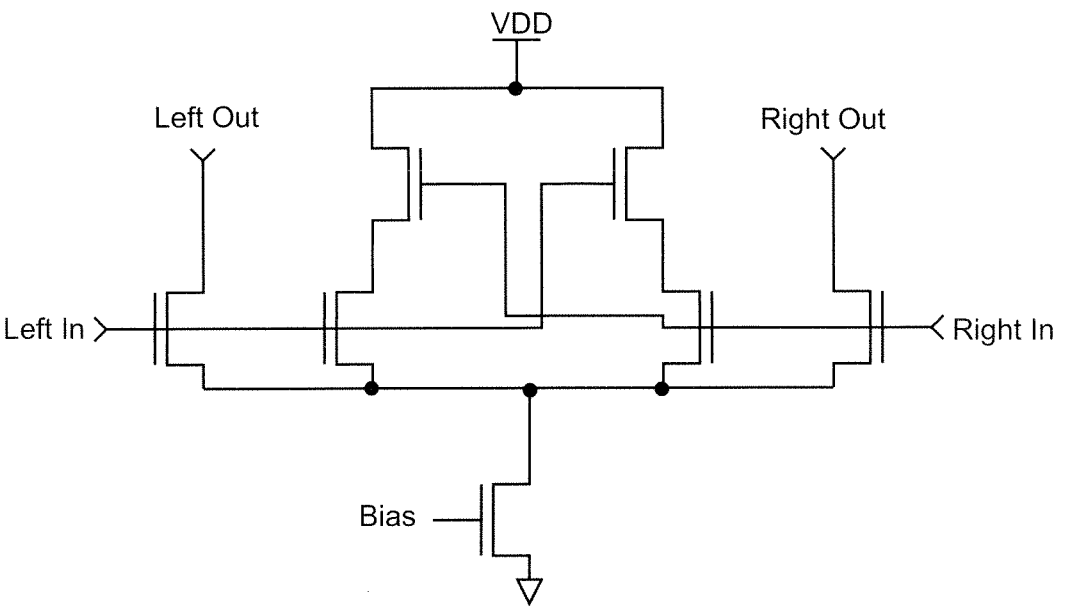

Figure 5.9: Schematic of the fully symmetric anti-bump circuit showing the connectivity used to achieve the desired non-similarity function.

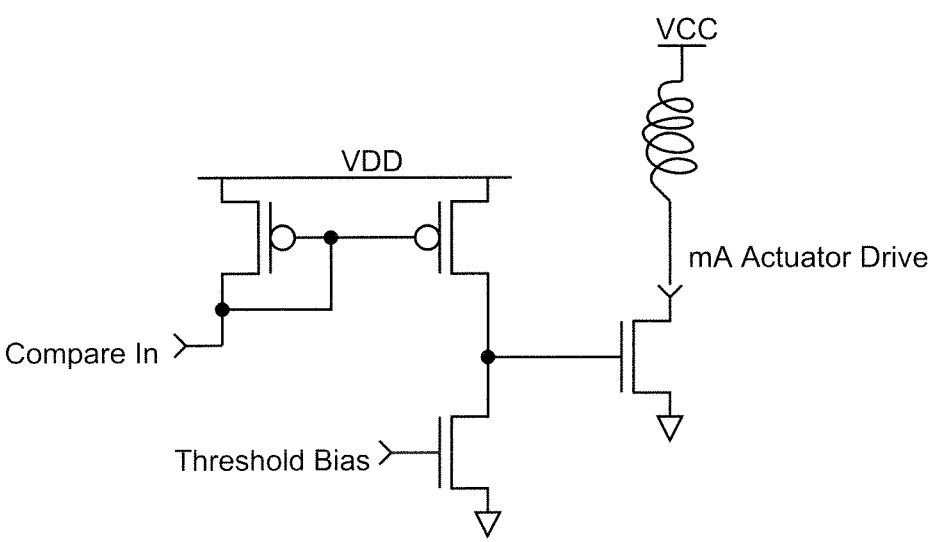

Figure 5.10: Schematic of the actuator driving circuit that shows the threshold circuitry along with the large output transistor necessary to sink the large currents needed to turn on the actuator.

tion for three inputs and one output is given by the following (assuming that the inputs to the system from the CT circuit are V_1, V_2 and V_3:

$$\begin{aligned} Output \;=\; & PostAmp(AntiBump(Filter(Buffer(PreAmp(V_0)), \\ & Buffer(PreAmp(V_1))), Filter(Buffer(PreAmp(V_2)), \\ & Buffer(PreAmp(V_2)))) \\ & + AntiBump(Filter(Buffer(PreAmp(V_2)), \\ & Buffer(PreAmp(V_3))), Filter(Buffer(PreAmp(V_3)), \\ & Buffer(PreAmp(V_4))))) \end{aligned} \tag{5.10}$$

where the $PostAmp()$, $AntiBump()$, $Filter()$, $Buffer()$ and $PreAmp()$ functions are represented earlier.

A die plot of the control/detection circuitry is shown in Figure 5.11.

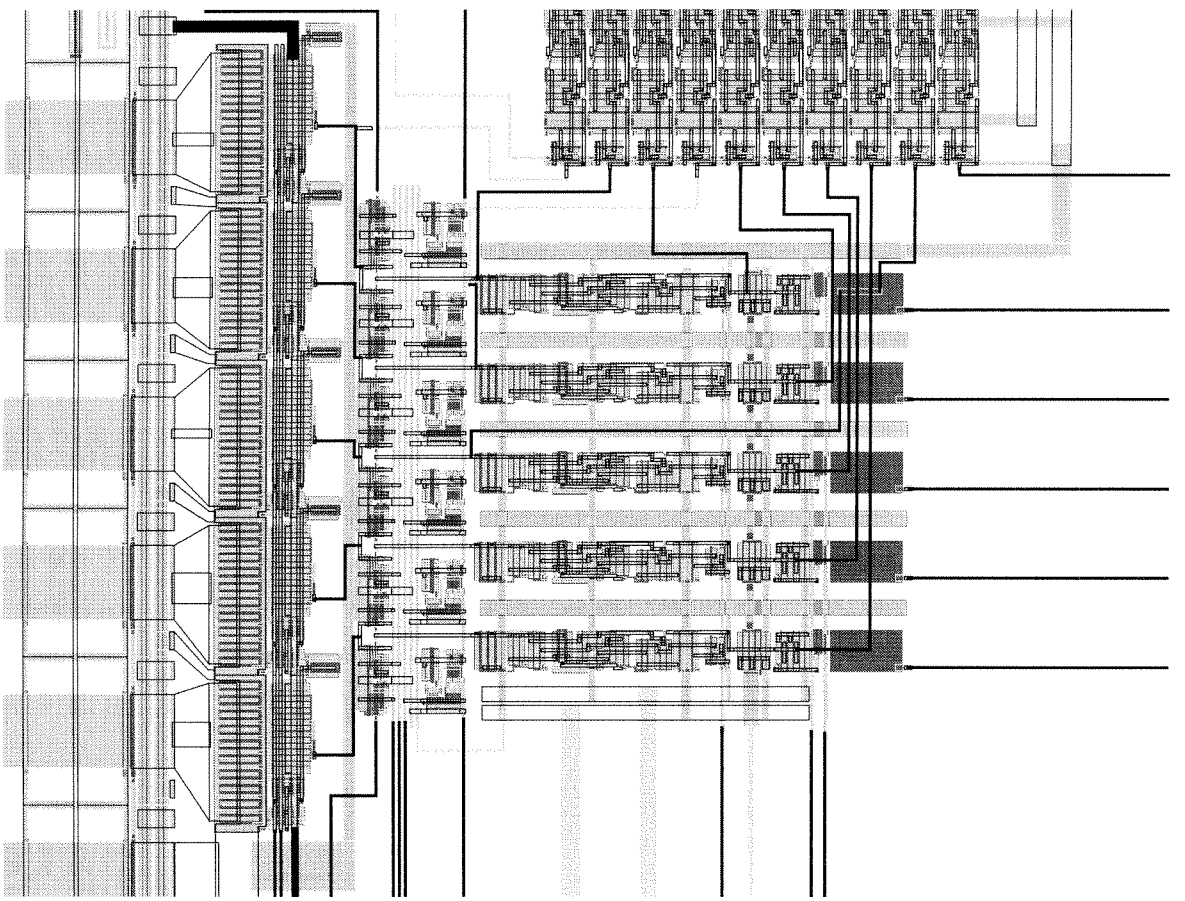

Figure 5.11: Dieplot of the control/detection chip. The information flows from right to left.

Chapter 6 Experiments and Results

In order to have the confidence to build a fully integrated system, we first need to verify the integrity of individual components of our system. We have conducted a variety of experiments that test the different components of the system. First, we have used signal generators to present synthetic signals to the detection/control chip to validate the functionality and performance of the design. We call this testing is the "lab bench" set of experiments. Second, operating in a laminar flow regime, mechanical macroscopic actuators generate known disturbances to verify the response of the system in the wind tunnel. We use this experiment to verify that we can connect different circuits of our system together. Third, an experiment is conducted where the system interacts with a fully turbulent channel flow in the wind tunnel to observe the response of the system. This experiment provides the data that gauges the usefulness of the high shear stress detection and control algorithm.

6.1 "Lab Bench" Results

The primary purpose of these experiments is to determine the basic operating parameters for the various circuit components by using known input signals. Measurement of the output allows us to determine quantities such as gain, bandwidth, and input range. These quantities are important to know to verify correct operation as well as to assist in setting the operating point for operation.

The first circuit to be verified is the pre-amplifier. The amplifier in Figure 5.4 has a gain–bandwidth product of approximately 500 kHz when the transistors are operated in the weak-inversion region. The transfer characteristic of the pre-amp at an arbitrary operating point is given in Figure 6.1.

The buffer (Figure 5.6) has a full power bandwidth of 1 MHz.

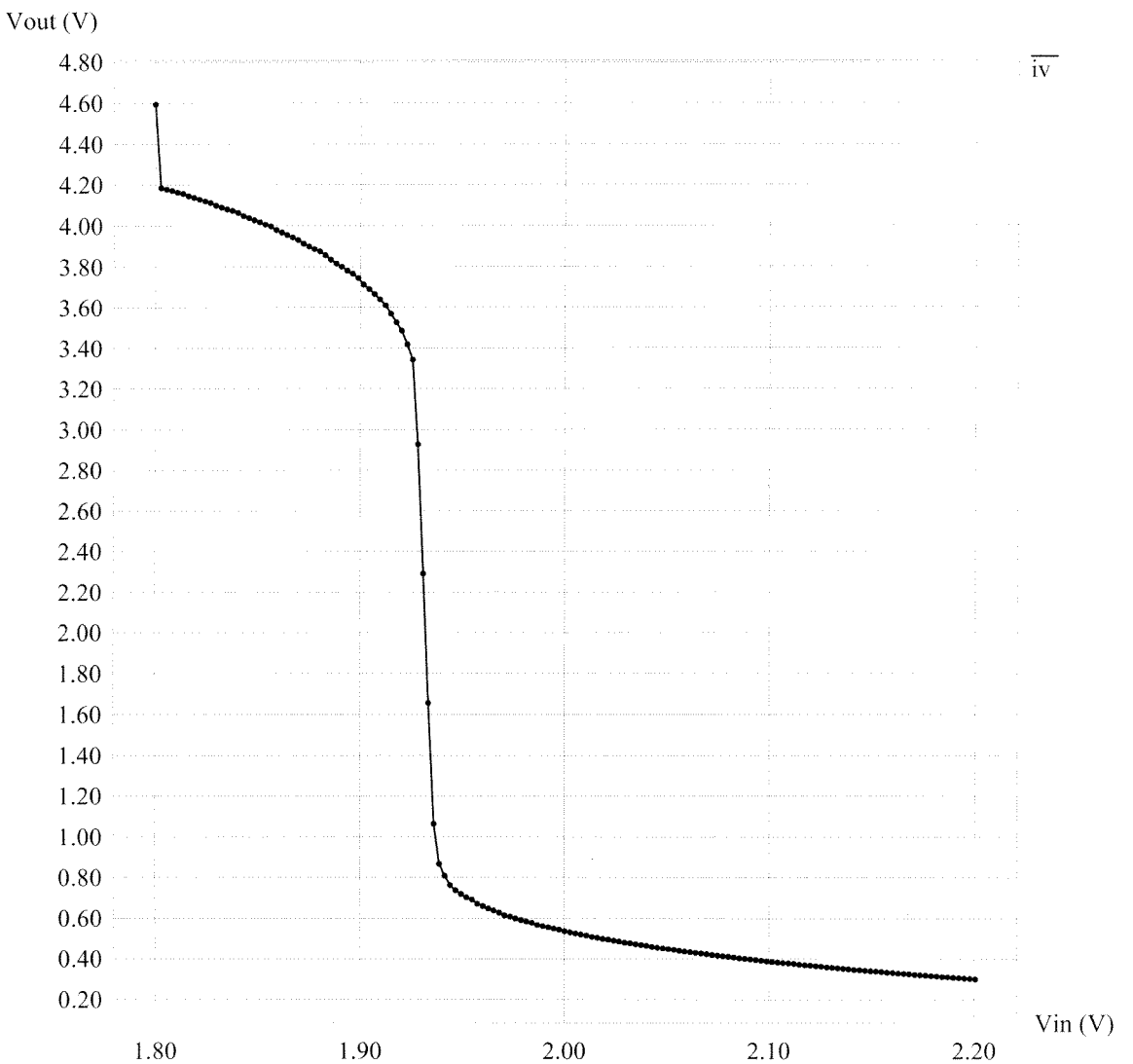

Figure 6.1: The voltage transfer characteristic of the pre-amplifier. The curve was taken at a frequency higher than the blocking frequency of the circuit.

We set the HRES bias to 0.5v, which corresponds to a small amount of spreading and low current levels before the adjoint cells "break." We have observed that if the bias is set too high, all of the interesting activity is "smoothed" away (i.e., the HRES resistors don't break). This results in our output driving signals always being low, i.e. the actuator is never turned on.

We set the anti-bump circuit bias to obtain about a 160mV dead zone. This is achieved in practice through a bias current ($V_{bias} = 0.80V$) sufficient to balance the dead zone width and the high shear stress detection capability. This means that ΔV between the inputs should be greater than 160 mV in order to get a sufficient current on the output. We have measured the current through the bump portion of the circuit (middle four transistors), plotted for different bias values in Figure 6.2. The currents that will flow through the anti-bump arms of the circuit are just a constant current (the maximum of the plotted current) minus the current plotted.

The actuator driver can sink approximately 30mA at a 1V drain to source (V_{ds}) drop. This corresponds to an impedance of about 30 Ω. Since this impedance closely matches the impedance of the surface-micromachined μ-actuator, we are close to an optimal power transfer at this operating point.

The overall latency of the processing system is about 40μs. Generally, the delay is fairly evenly spread across the different system blocks. However, the HRES circuit does contribute a bit more delay than the other circuits because of its low bias voltage. This latency is acceptable for the system to detect the high shear stress regions.

6.2 Experiment in the Laminar Flow Regime

This experiment consisted of generating a periodic disturbance in a laminar flow in order to check the response of the signal processing and actuator driving circuitry. A macro-sized actuator (eliptical - 1mm x 4mm) attached to a solenoid driven by an external signal generator created the flow disturbance. The actuator disturbed the flow to create a vortex pair like entity. This disturbance is akin to generating a

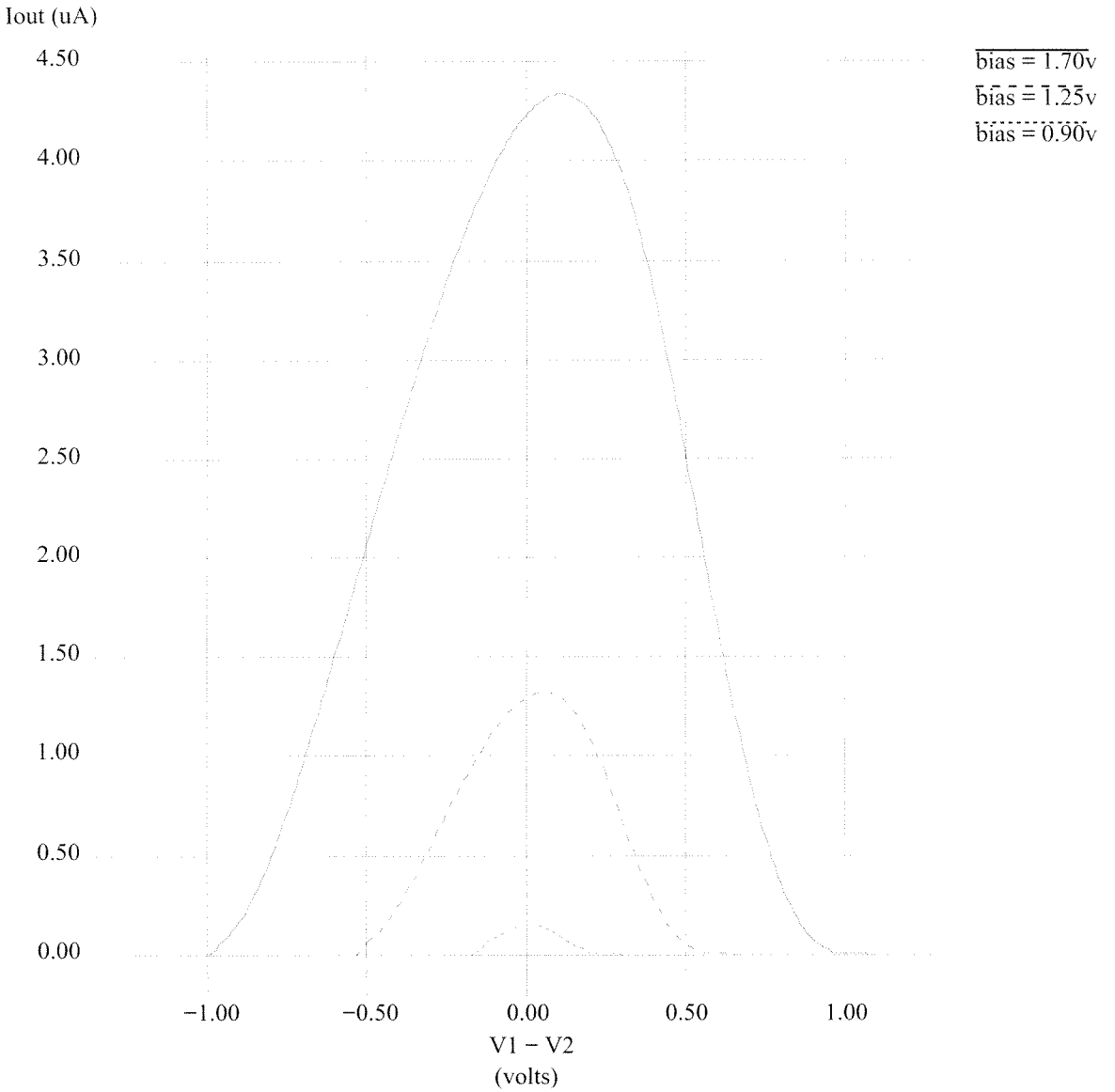

Figure 6.2: The current as measured through the middle leg of the anti-bump circuit. The current plotted is called the "bump" current as opposed to the anti-bump current that we actually use. The width of the "bump" corresponds directly to the dead zone size. Measurement of the bump current does not interfere with the operation of the circuit in the system.

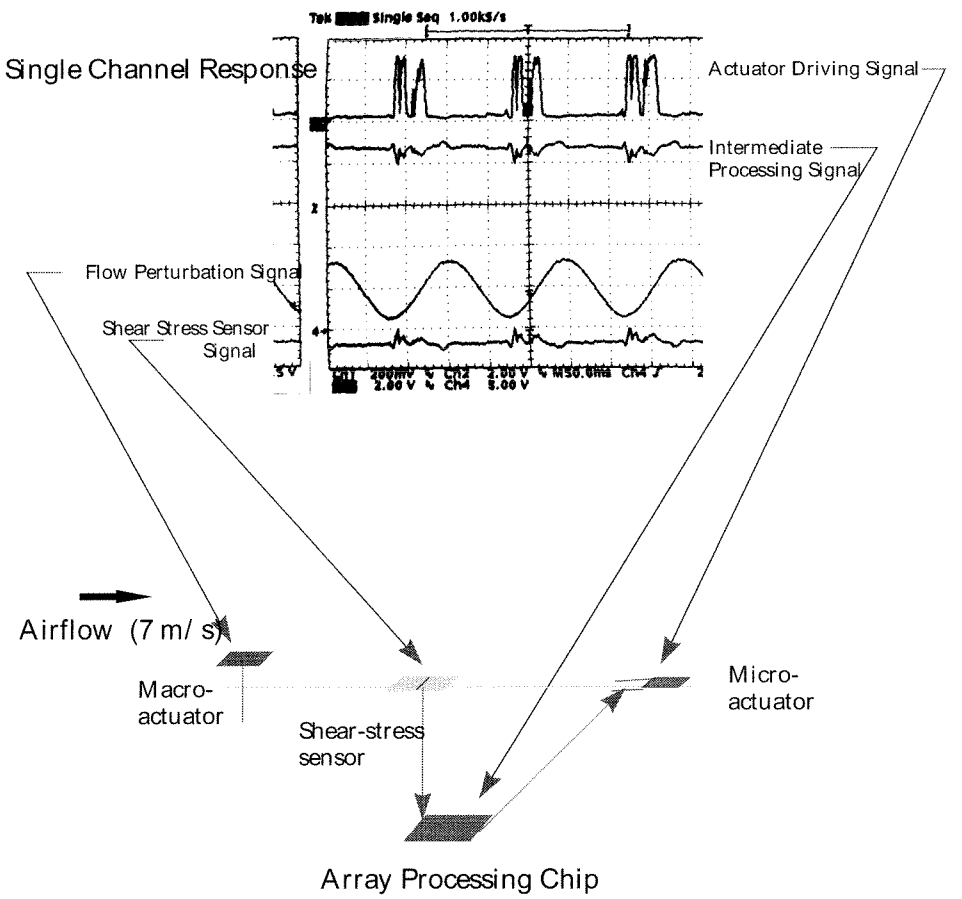

Figure 6.3: Diagram of the setup for the laminar test.

vortex pair in the viscous sublayer of the turbulent boundary layer. The sensors in this experiment were controlled by a constant current circuit. From Section 4.2.1, we know that the dominant time constant of this system is the thermal time constant of the sensor ($\approx$ 300 Hz) which is too low for the turbulent experiments. We chose a perturbation frequency of about 7 hertz. This low frequency allowed for adequate controllability of the solenoid. We depict this experiment in Figure 6.3.

6.2.1 Actuator Efficacy

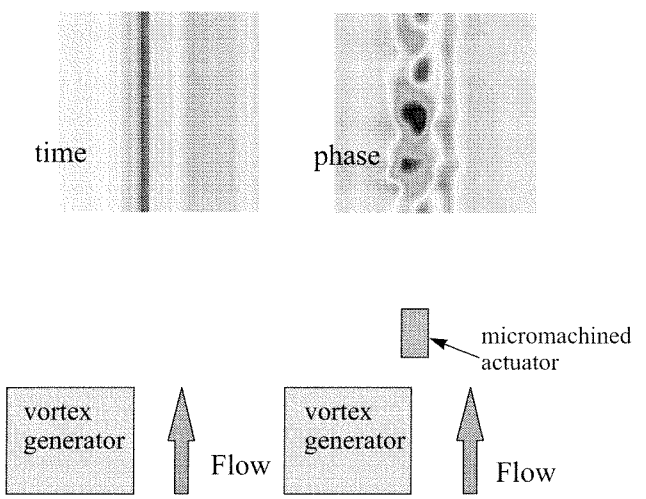

Figure 6.4: Diagram showing the test setup as well as the contour plot results for the actuator test. The purpose of the experiment was to gauge the effectiveness of the actuator on diminishing the effect of a static vortex pair. Figure courtesy of Dr. Steve Tung.

Another experiment was concurrently conducted by Dr. Steve Tung to ensure that the μ-actuators can affect the flow and *potentially* reduce drag. This experiment was conducted (depicted in Figure 6.4) in the laminar flow regime with an artificially generated static vortex pair. Using a periodic driving signal on the μ-actuators, a 10% reduction in the averaged shear stress was measured. This result does not include the consideration of form drag which would reduce the effectiveness of the μ-actuator and hence the overall drag. The key point to note from this experiment is that the shear stress caused by a vortex pair can be reduced by the motion of an μ-actuator. If a similar methodology can be employed in the turbulent regime

using the μ-actuators when a region of high shear stress is detected, we may be able to accomplish similar reduction of the shear stress and ultimately drag.

6.3 Turbulent Flow Testing

The design goal of the system is to mitigate regions of high shear stress in our experimental setup with a fully turbulent airflow profile. The turbulent profile is selected by moving the site of the experiment downstream in the wind tunnel. From Section 1.7, this site is located approximately two-thirds of the distance along the wind tunnel. The Reynolds number used in this experiment is approximately 10^4 and is varied by adjusting U_c (see Section 1.3). The centerline velocity of our wind-tunnel channel varies between 8 $\frac{\text{m}}{\text{s}}$ and 20 $\frac{\text{m}}{\text{s}}$ (28 km/hr and 72 km/hr).

For this experiment, we mount a μ-sensor array in a Plexiglas plug. The plug also houses three μ-actuators mounted further downstream. The number of μ-actuators is limited by physical size constraints of the μ-actuator. Five outputs of the CT shear stress sensors are fed into the detection/control chip. Since there are fewer actuators available, we only use three outputs of the detection/control chip to drive the available actuators.

We sample the data (from both the CT shear stress sensors and the control/detection circuit) at a rate of 5kHz. This sampling rate is then converted to a distance based on the centerline flow velocity by using the formula:

$$\begin{aligned} D &= \frac{\psi\, U_c}{5\,\text{kHz}} \\ &= 1.2\,10^{-4}\, U_c \end{aligned} \tag{6.1}$$

where $\psi = 0.6$ is an empirically derived constant that relates the wall convection velocity to U_c.

The combined sets of data can be analyzed to see the potential for high shear stress detection and eventually drag reduction. Figure 6.5 graphs the single col-

umn response of both shear stress sensor output and the corresponding channel from the detection/control chip. The data shows that the circuit responds to the high values of shear stress. The input signal has been processed correctly because the output does not respond to the small changes in the signal only the larger ones.

6.3.1 Shear Stress Imaging

A two-dimensional flow picture of instantaneous surface shear stress is shown in Figure 6.6. From left to right, we plot the full span (25 sensor) recording, an enlargement of the middle three sensor responses and the outputs of the detection/control chip corresponding to those three sensor inputs. The data was recorded in a turbulent flow regime with a free stream velocity of 10 $\frac{\mathrm{m}}{\mathrm{s}}$. By sampling the spanwise dimension in time, we can construct a two-dimensional map indexed by the sensor position and time. We can convert the time index into a distance by the use of Equation 6.1. We also repeated the experiment at a higher flow velocity (Figure 6.7). Notice that the size of the high shear stress region is smaller on each higher velocity subplot of Figure 6.6 and Figure 6.7 which means that the size of the vortex pairs has decreased as expected.

6.3.2 Detection/Control Output

In Figure 6.8, Figure 6.11, and Figure 6.14 one can see that the distribution of input shear stress values has changed as the velocity has increased. The maximum value of shear stress τ_w has gone from 0.45 to 4.5. The distribution also has a greater tail towards higher shear stress values.

Figure 6.9 shows the transfer curve for three channels of the detection/control chip from data recorded at 12 $\frac{\mathrm{m}}{\mathrm{s}}$. The data indicates that for large values of shear stress, the detection and control circuitry would, in fact, be large enough to turn on an actuator (the criteria being voltage greater than about 1.5v). Figure 6.12 shows the same plot with the data recorded at 15 $\frac{\mathrm{m}}{\mathrm{s}}$. Figure 6.15 plots the response at the fastest speed tested of 20 $\frac{\mathrm{m}}{\mathrm{s}}$.

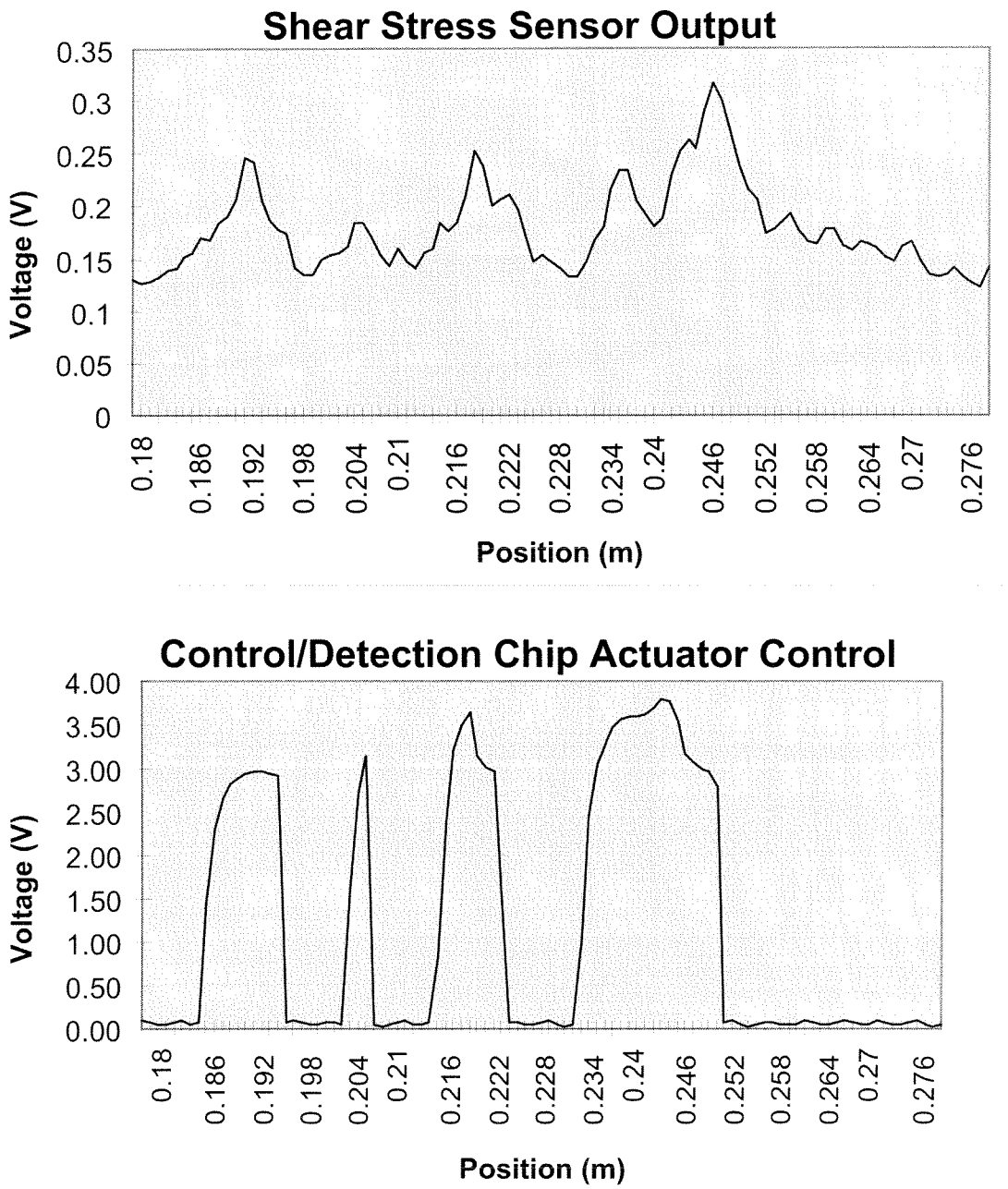

Figure 6.5: Strip chart of the shear stress CT voltage and the output of the control/detection circuitry. The x-axis of the plot has been converted to distance from time by the use of Equation 6.1.

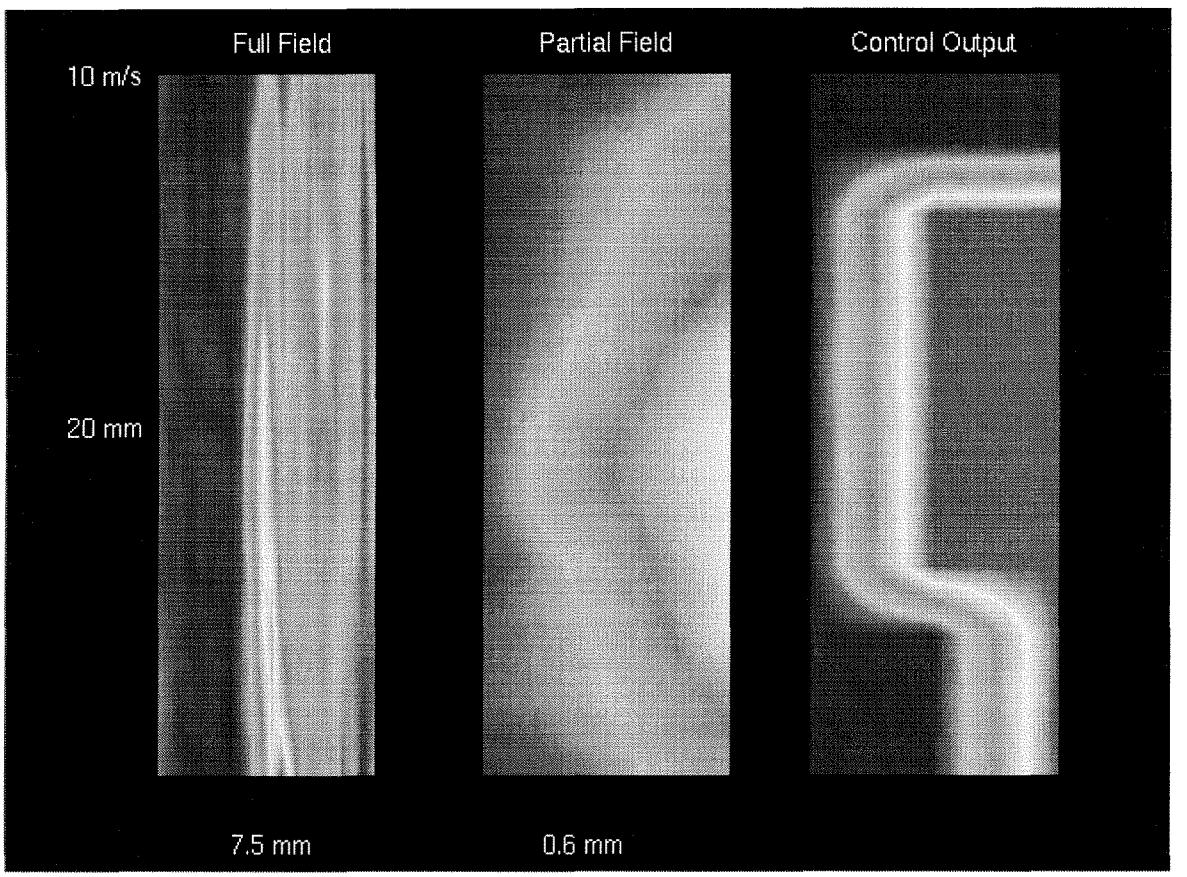

Figure 6.6: Picture of the turbulent shear stress imaging along with control outputs recorded at 10 $\frac{\mathrm{m}}{\mathrm{s}}$.

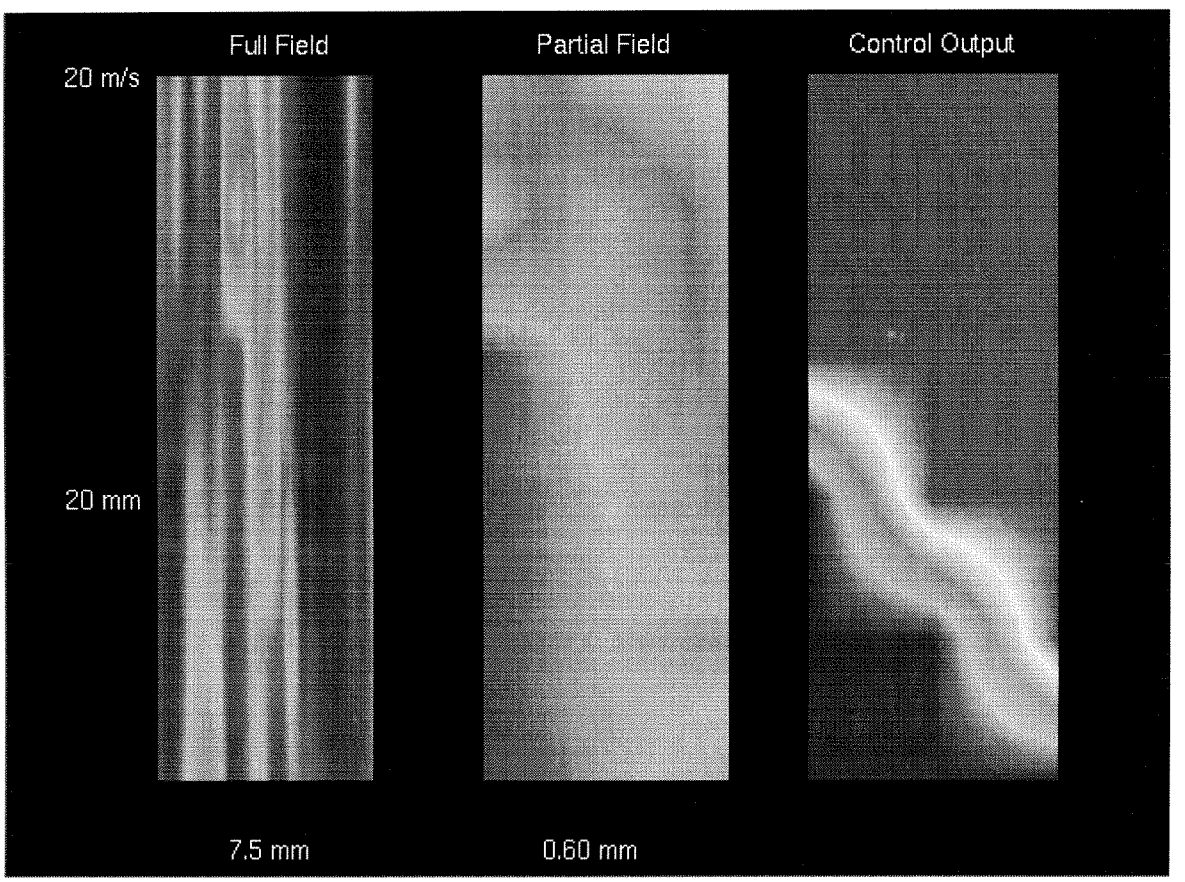

Figure 6.7: Picture of the turbulent shear stress imaging along with control outputs recorded at 20 $\frac{\mathrm{m}}{\mathrm{s}}$.

The distribution of the output values mirrors the input distribution. Figure 6.10 has a similar distribution as Figure 6.8. Similarly, Figure 6.13 mirrors Figure 6.11. Figure 6.16 does deviate from Figure 6.14 most likely due to a misadjusted threshold setting.

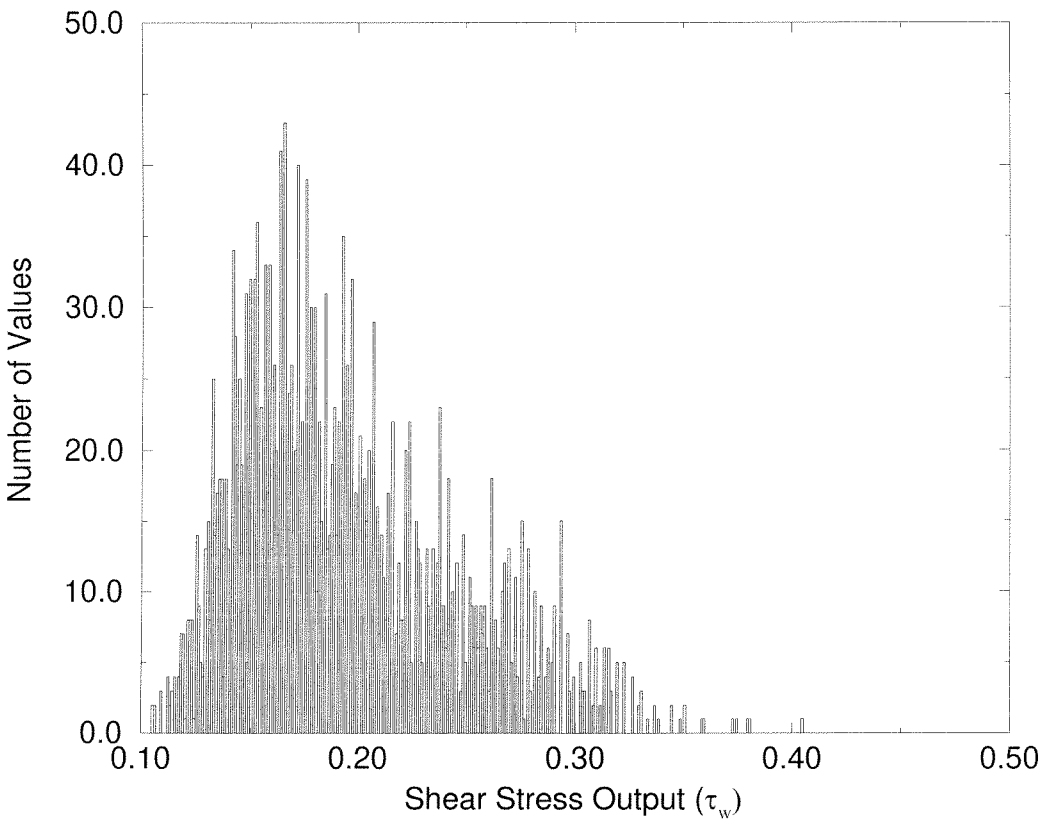

Figure 6.8: Distribution plot of shear stress values at a low speed (12 $\frac{m}{s}$).

Drag Reduction

In these experiments, the μ-actuators do not yet have the mechanical frequency response or stiffness to follow the actuation signals from our chip. Thus, only a very rough estimate of the drag reduction that our system could potentially provide is possible. We assume that the μ-actuators are about 10% effective in mitigating the high shear stress based on the experiments using the μ-actuator to affect a static

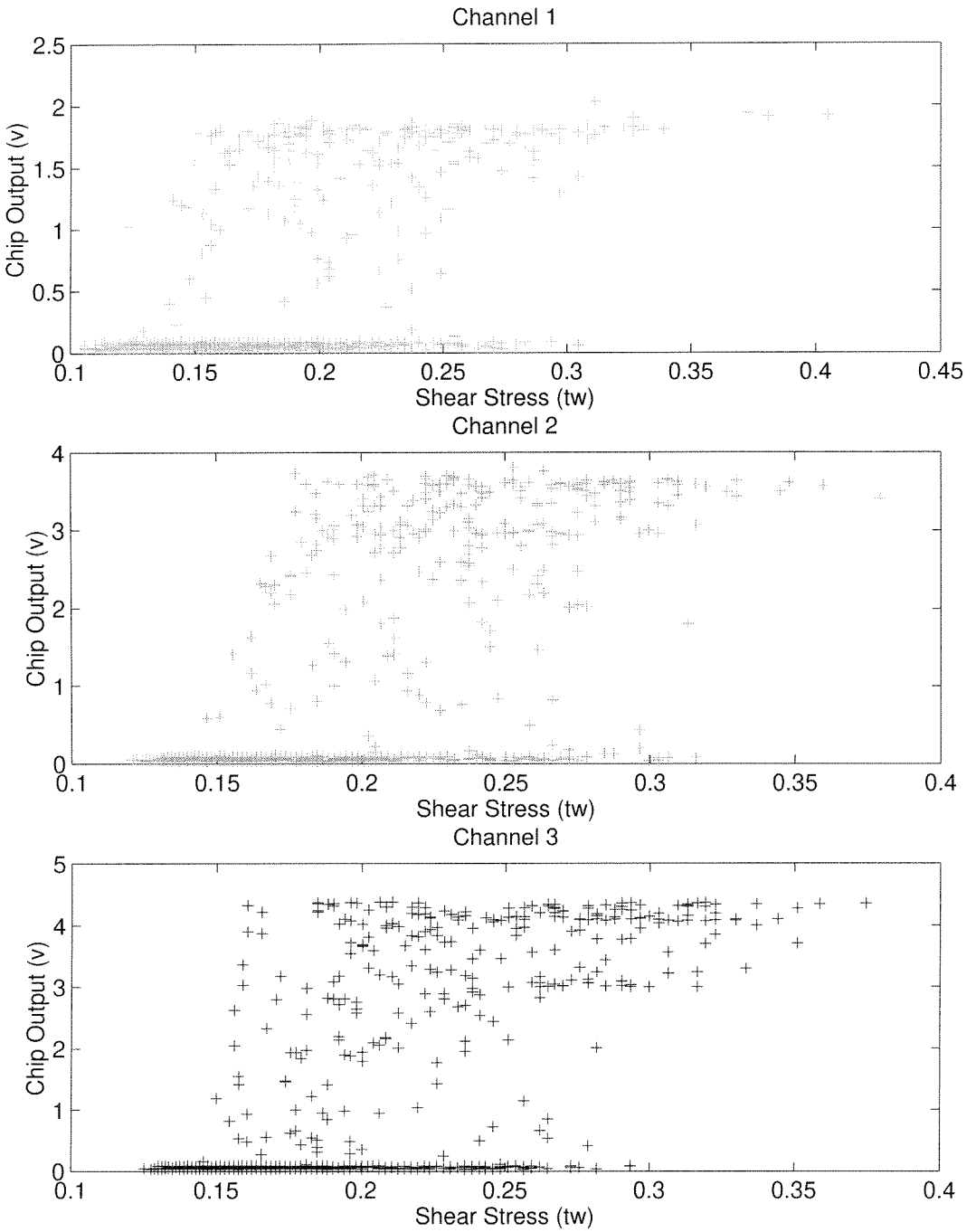

Figure 6.9: Transfer curve between input shear stress voltage and output actuation at low speed (12 $\frac{\mathrm{m}}{\mathrm{s}}$).

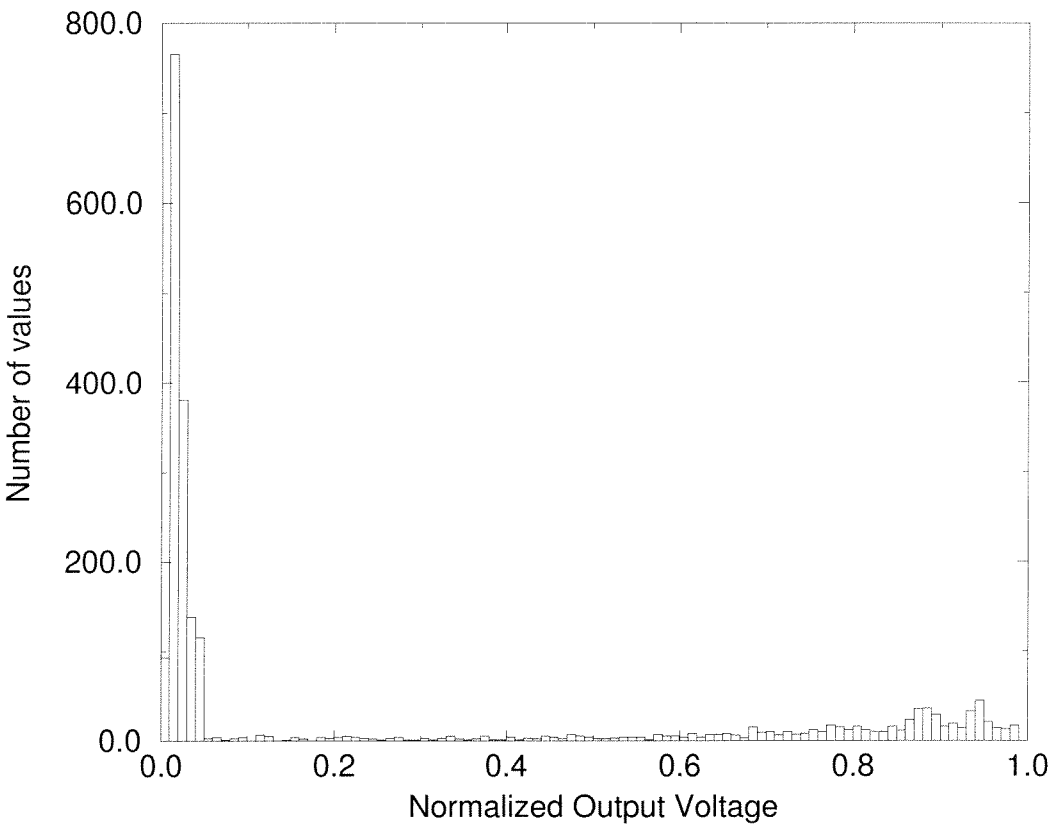

Figure 6.10: Distribution plot of output control signal at a low speed.

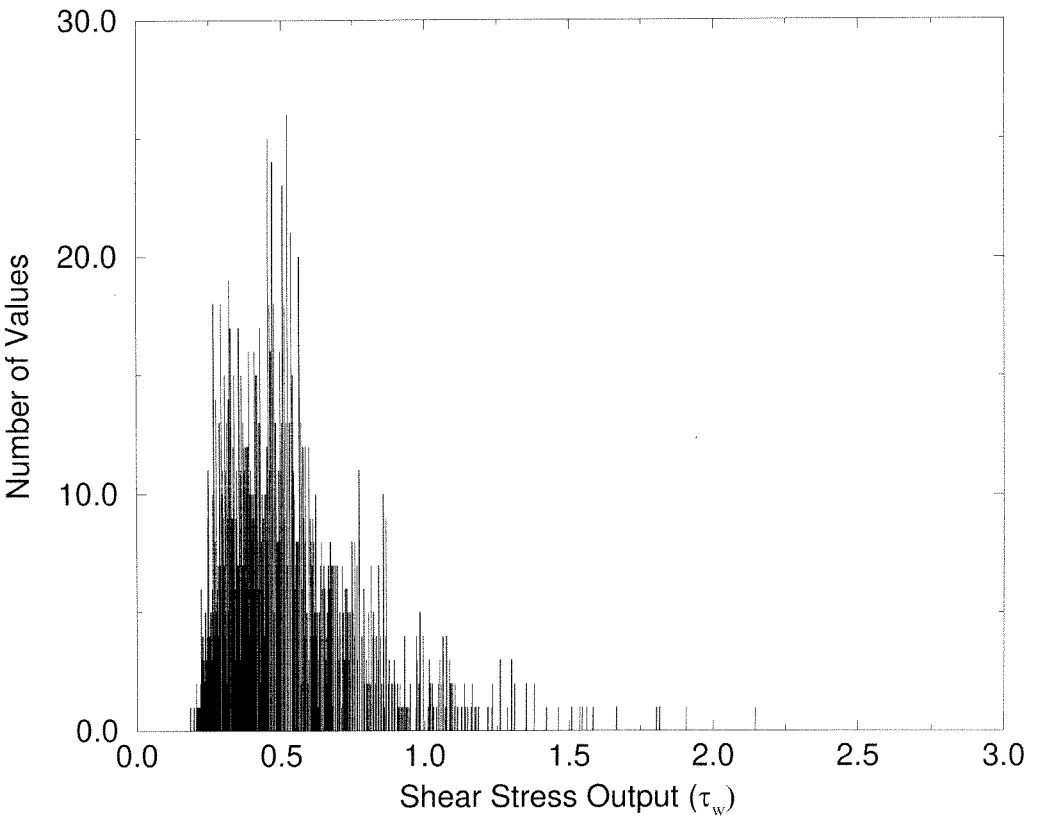

Figure 6.11: Distribution plot of shear stress values at a medium speed (15 $\frac{m}{s}$).

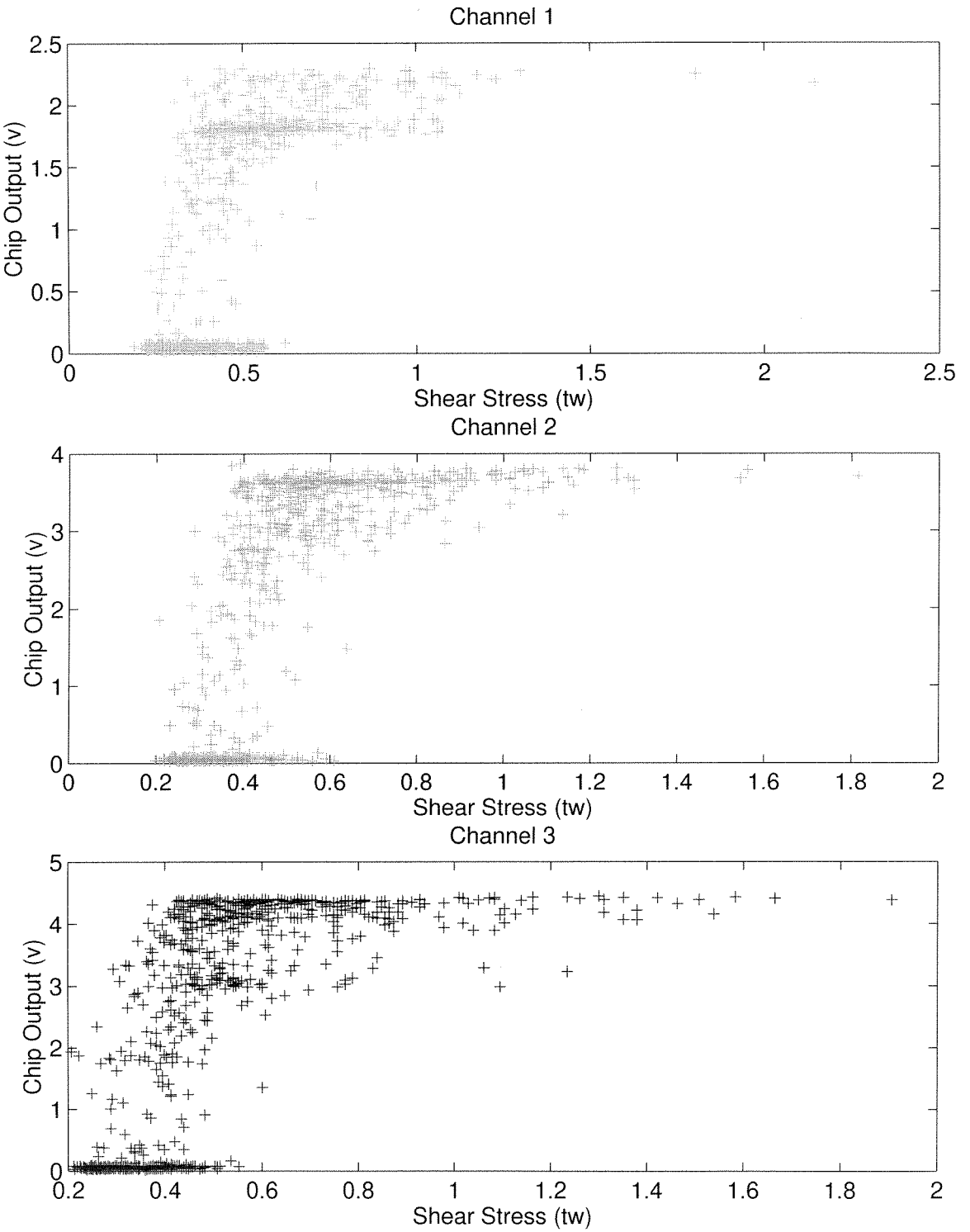

Figure 6.12: Transfer curve between the input shear stress voltage and output actuation at a medium speed (15 $\frac{\text{m}}{\text{s}}$).

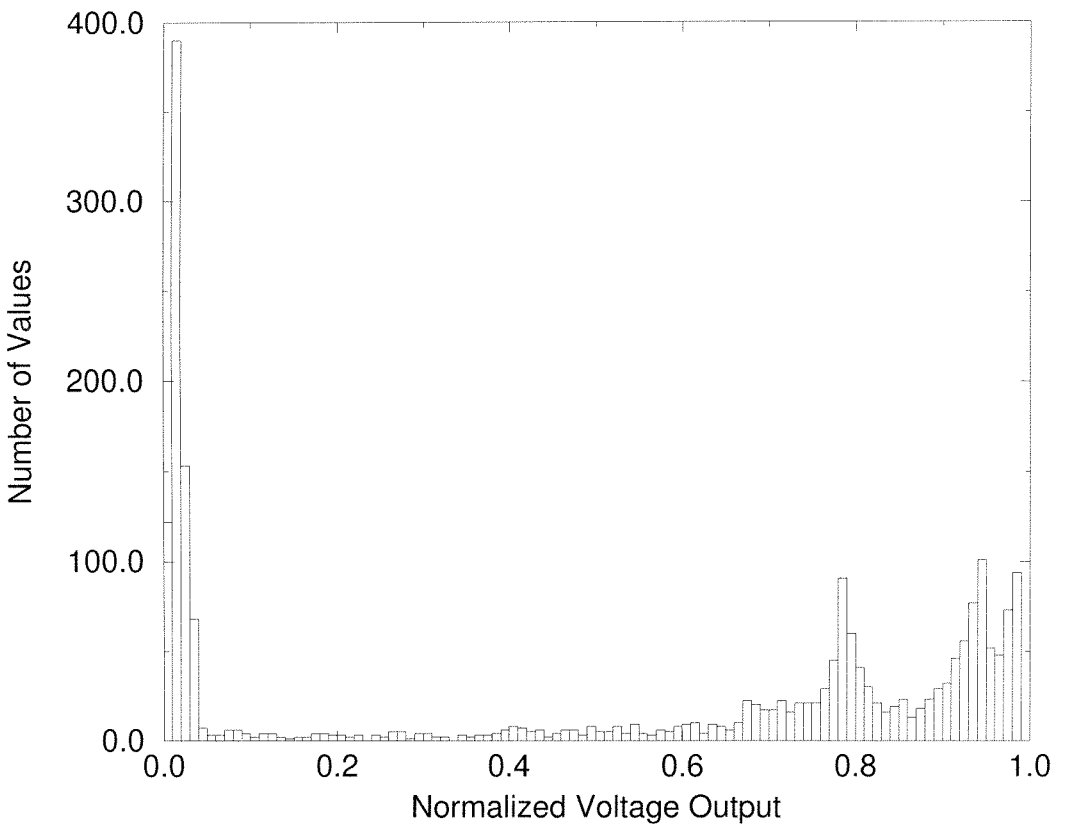

Figure 6.13: Distribution plot of the output control signal at a medium speed (15 $\frac{\mathrm{m}}{\mathrm{s}}$).

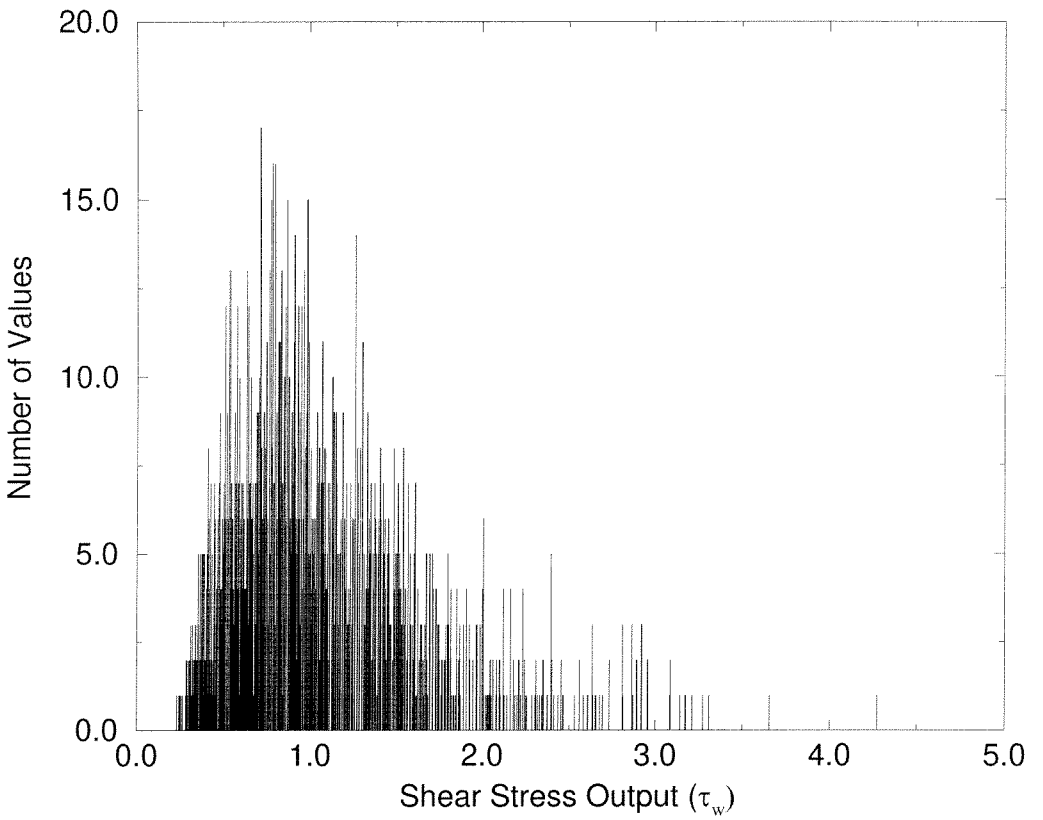

Figure 6.14: Distribution plot of shear stress voltage values at a high speed (20 $\frac{\mathrm{m}}{\mathrm{s}}$).

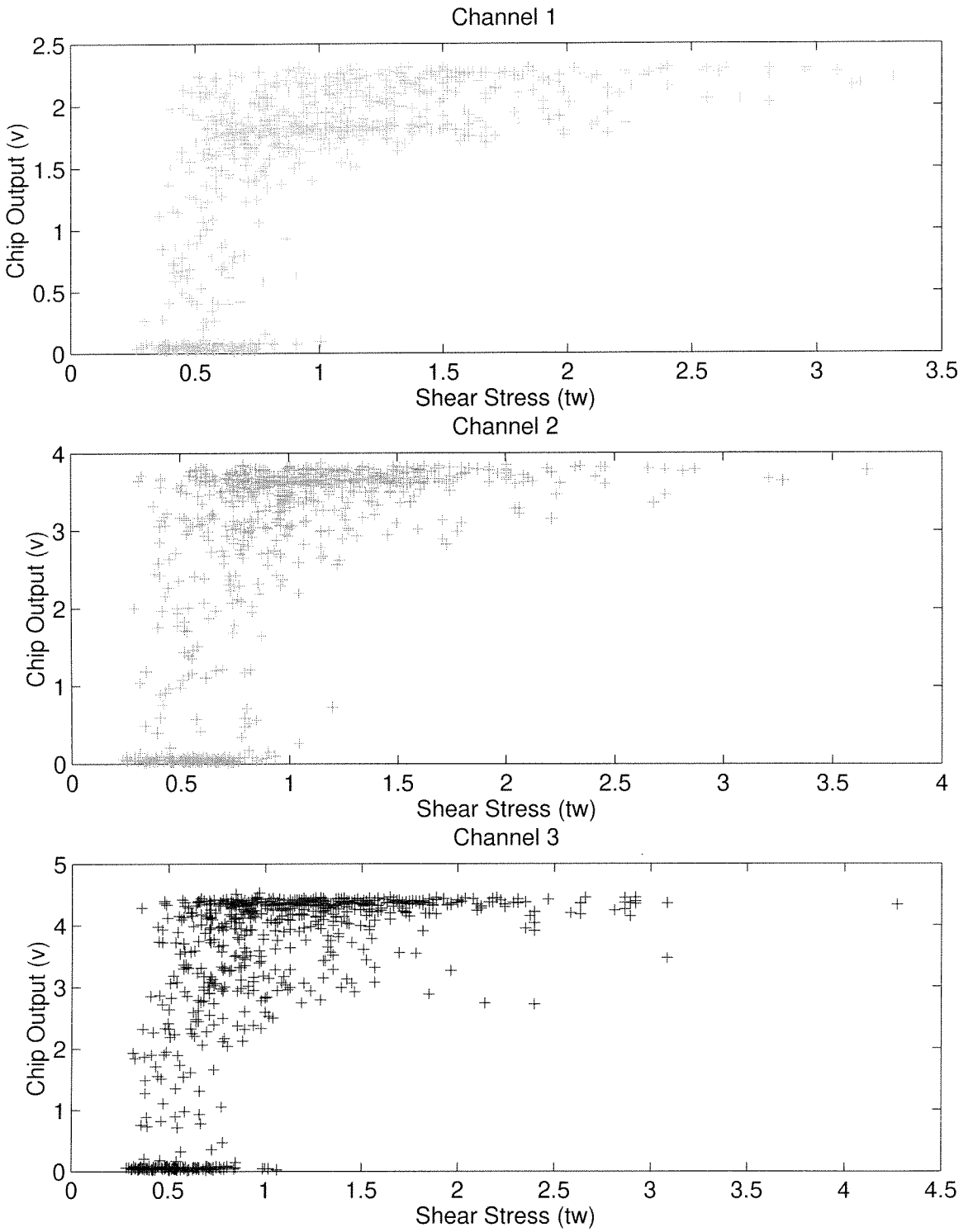

Figure 6.15: Transfer curve between the input shear stress voltage and output actuation at a high speed (20 $\frac{\mathrm{m}}{\mathrm{s}}$).

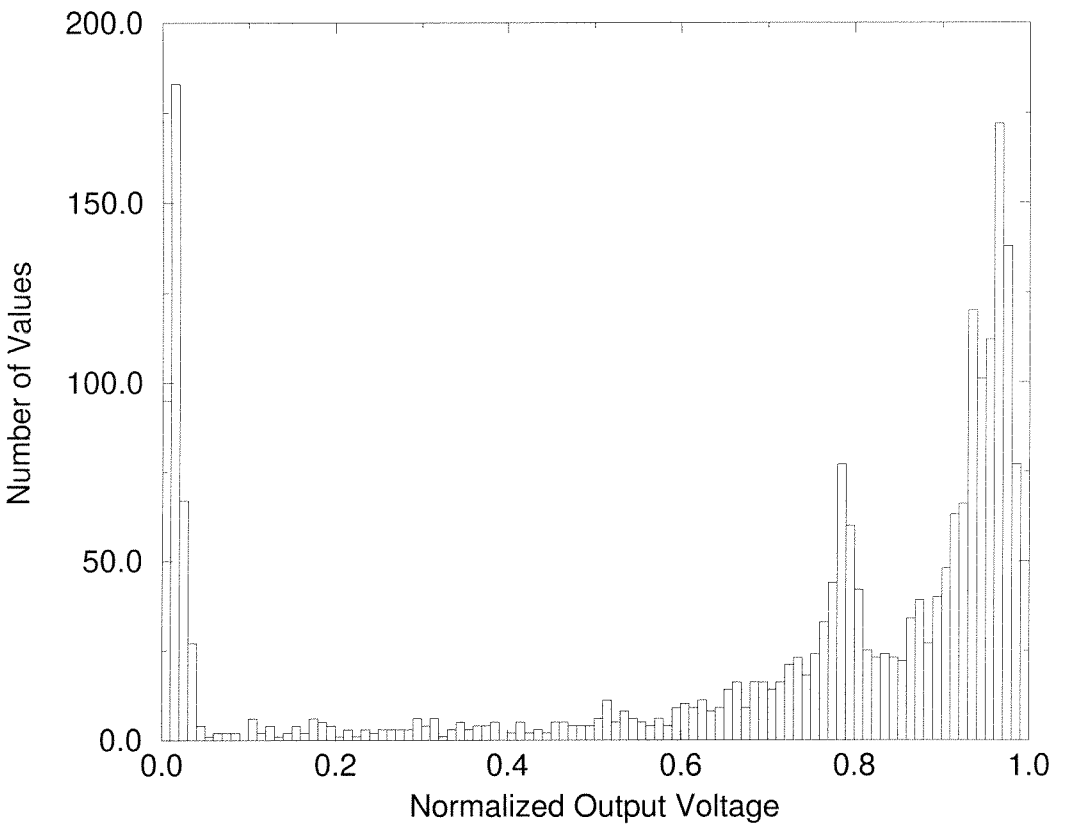

Figure 6.16: Distribution plot of the output control signal at a high speed (20 $\frac{\text{m}}{\text{s}}$).

vortex pair (Section 6.2.1). While the conditions of the flow are substantially different, these experiments serve as the best evidence available of the actuator efficacy. The calculation involves counting the number of high shear stress values at which the detection/control circuit responded and calculating the reduction in total drag if those high values were 10% smaller. From our recorded data, we estimate the drag reduction is approximately 1%.

Part III

Integrated System

Chapter 7 Integrated System

7.1 Background

The final goal of this work is to build an integrated system of micromachining and CMOS circuitry. While this general goal has been accomplished by many other researchers, our effort is to specifically integrate specialized MEMS structures onto a standard CMOS process. This type of integration we wish to perform presents us with many hurdles to overcome. Many micromachining systems include some circuitry on the same die, but it is generally not as sophisticated as the process one might find on a CMOS standard process. Alternatively, if the substrate starts out in a CMOS-only fab, the types of micromachining components that can be constructed are of a limited variety. We need the best of both worlds, i.e. a reasonably good CMOS process plus complex micromachining capability.

7.2 TinyChip Integration

The MOSIS service's 2.0 μm TinyChip fabrication capability provided to be an attractive method to investigate what issues may arise in performing a complete integration process. The most important concern was whether or not the CMOS transistors would retain their characteristics after post-processing. In particular, the most important characteristics for analog circuitry include the threshold voltage, the flat-band voltage, carrier mobility, channel doping and junction leakage currents.

7.2.1 Compatibility of CMOS with MEMS

To answer the question of process compatibility, we fashioned several different chips. The first few series of these chips were learning exercises because the chips were rendered useless by the standard CMOS processing. The complex interaction of etch layers when there is no material underneath to be etched caused severe amounts of over-etching into supposedly protected layers. The basic concern of compatibility is the high temperatures necessary for the post-processing. The high temperatures preclude the use of aluminum (for the metalization layers) before the high-temperature post-processing steps since it would melt if deposited. This requirement meant that all the fabricated CMOS structures had to be connected without metal. Since the Orbit/MOSIS 2.0 μm process is not a buried contact[1] process, only solitary transistors can be reasonably fabricated for subsequent testing of the current versus voltage (IV) characteristics of the transistors. If no significant changes are measured, then the idea of integration can proceed.

The chip design which yielded the consistent results was fabricated without extra etch layers. Photoresist was then painted on with a fine brush to protect the transistors. Once hardened, the chip was etched in hydrofluoric acid (HF) to remove the oxide to allow direct probing of the drain, source, gate and bulk contacts. The drain, source and bulk contacts were regions of diffusion or well, and the gate contact was a region of polysilicon.

The results of this experiment demonstrate that there was no measurable change in the transistor characteristics even in the sensitive subthreshold (weak-inversion) region. We chose the smallest allowable transistors as our representative sample because they are potentially most sensitive to temperature induced problems. The plots of the current versus voltage curves are shown in Figure 7.1 and in Figure 7.2. The approximate threshold voltage is about 0.85v (NMOS). The difference between the before and after curves is well within the measurement error (approximately

[1] A buried contact process is one in which polysilicon (typically used as gate material) can also make a ohmic contact to a source or drain.

5%). The most important result is that the subthreshold slope[2] did not appreciably change. This is important because many of the circuit parameters like gain, offset, and bandwidth are sensitive to the subthreshold slope.

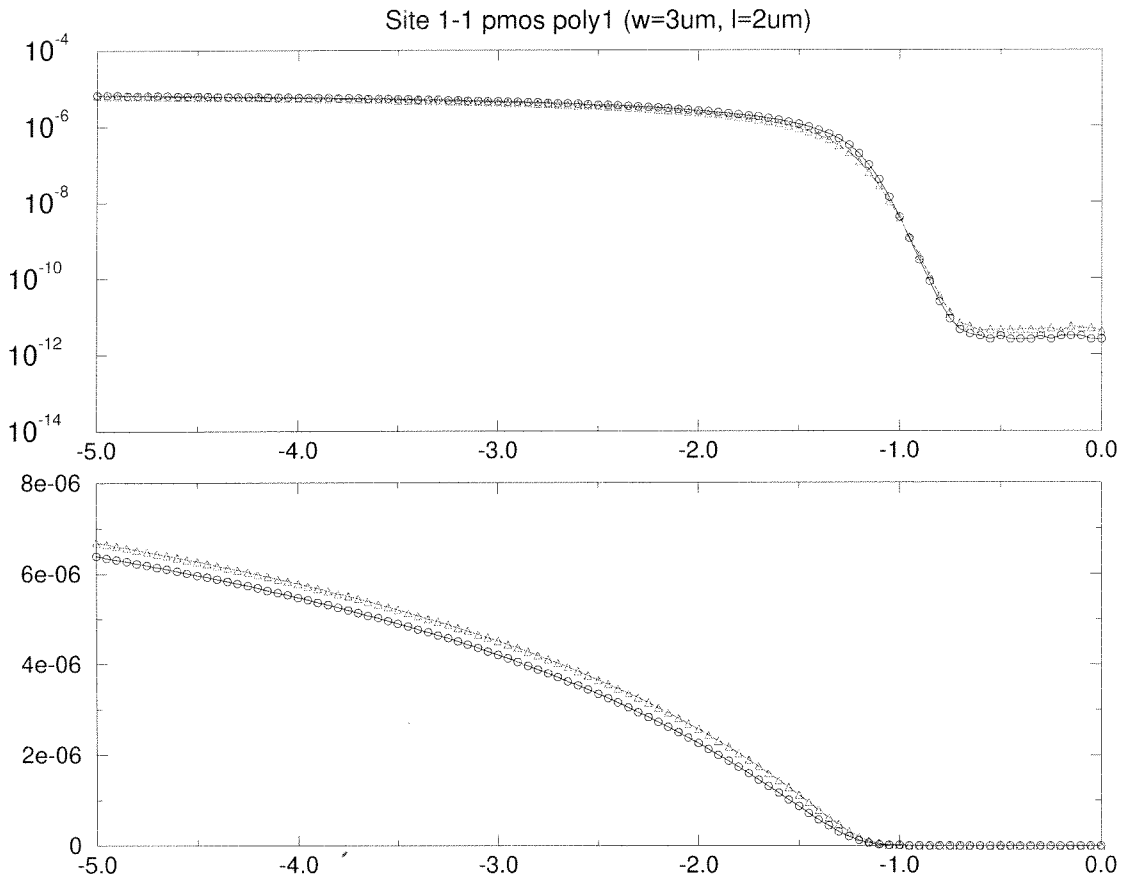

Figure 7.1: PMOS transistor characteristics before and after high temperature bake which shows a shift of less than 5% in the current versus voltage plot.

7.2.2 Integrated CMOS Sensors

An attempt was made to fabricate the sensors using the layers available in a standard CMOS process. Figure 7.3 shows a plot of the sensors at the edge of the chip. Gate polysilicon was used as the sensor material. The post-processing involved etching the silicon dioxide (oxide) from around the sensor beam to allow

[2] Also known as $\frac{1}{\kappa}$.

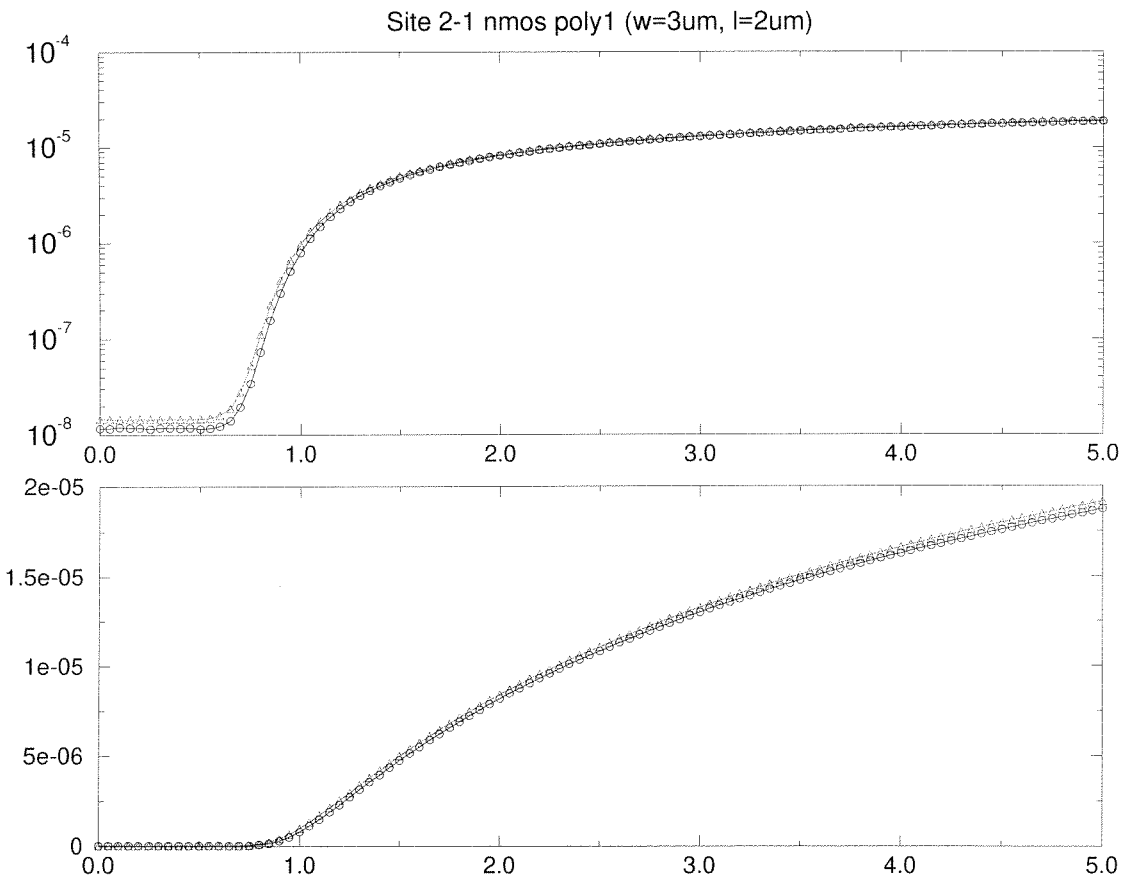

Figure 7.2: NMOS transistor characteristics before and after high temperature bake which shows less than a 2% shift in the current versus voltage plot.

heat transfer to the air. In this attempt, a buffered HF solution[3] was used as the etchant. Because of stiction problems[4], a CO_2 drying was also attempted after the release step. This procedure allowed the MEMS structures to go directly from being immersed in a liquid to being free of any immersant through the evaporation of the liquid CO_2. Unfortunately, this experiment did not yield any functional sensors because the etching was too imprecise. It proved too difficult on a small die to balance the etching necessary for the sensors while preventing over-etching on the circuitry. Hence the conclusion is that MEMS structures on small (2.1 mm by 2.1 mm) dies are difficult to work with reliably. If this type integration is necessary, one has to use a whole wafer for acceptable results.

7.2.3 Integrated CMOS Actuators

We also attempted to fabricate parts necessary for a μ-actuator on a CMOS die. Basically, a metallic coil was built to create a magnetic field to lift a permalloy coated flap placed on top of the chip. The coil (Figure 7.3) was fabricated. However, subsequent testing failed to demonstrate any motion of the overlaid flap when current was run through the coil.

7.3 First Wafer-Scale Integration

Many changes were made to the system to facilitate integration. The designs of the sensors, μ-actuators and the layout of the electronics were all modified to allow the systems to be combined on a common substrate.

The first attempt at wafer scale integration involved fabricating the design at the University of California at Berkeley's MICROLAB. The fabrication facility produces a 2.0 μm double polysilicon CMOS process. Because of the relative similarity of the Berkeley and MOSIS/Orbit processes, both could be used without a great deal of extra layout effort. However, because of questionable transistor yield on

[3] Also known as pad-etch.

[4] A phenomena where MEMS parts stick together because of a variety of forces

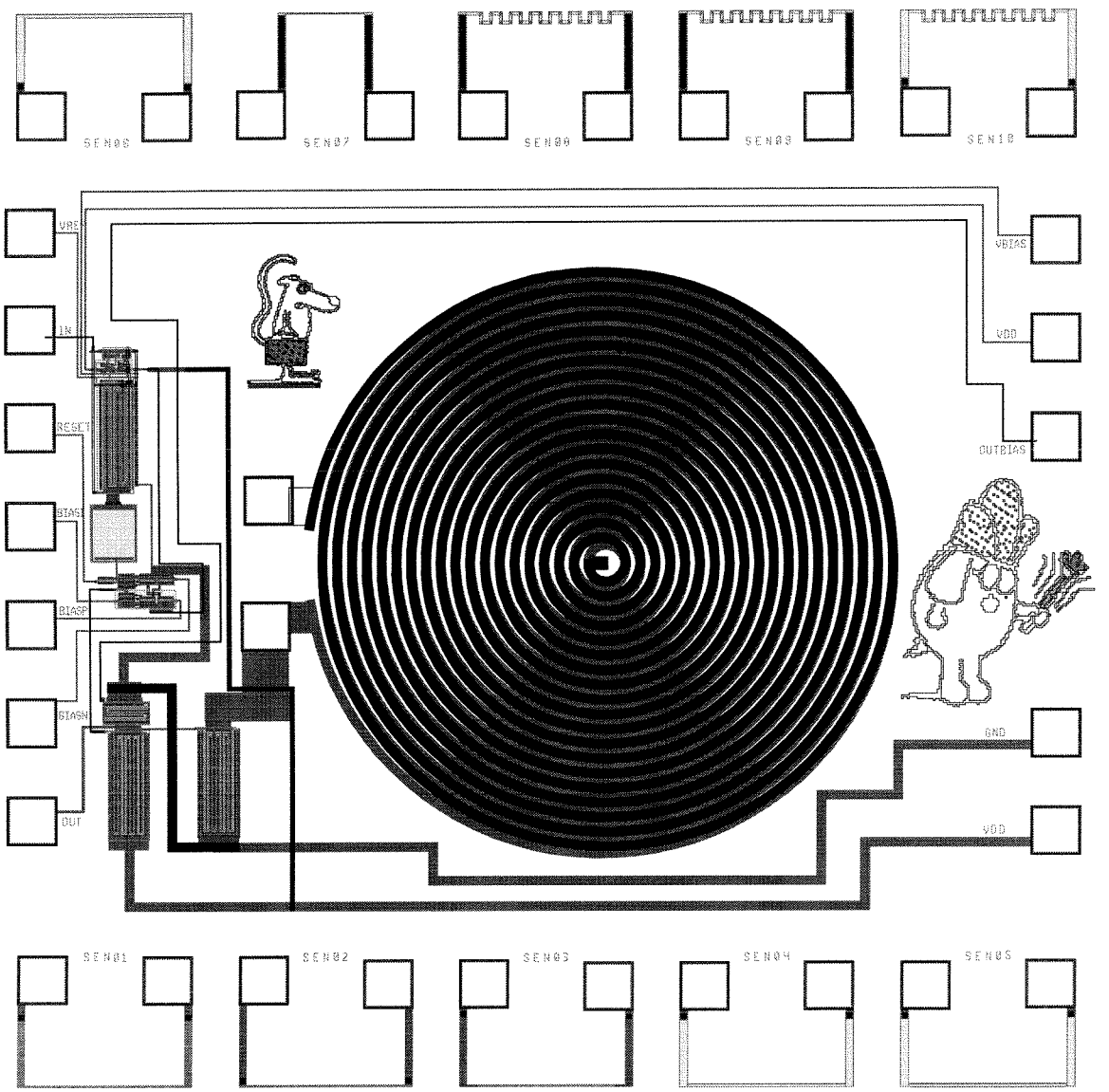

Figure 7.3: Dieplot of a metal coil structure using two CMOS metal layers.

the Berkeley line, many changes were finally incorporated into the layout (especially the electronics). For instance, all the contacts were fabricated to be 5 μm by 5 μm which is over a 600% increase in contact area (from the standard 2 μm by 2 μm).

7.3.1 Modified Low Voltage Sensor (Type II)

The entire CT circuit is intended to be built on a single substrate. We have demonstrated the efficacy of building an on-chip opamp capable of driving the resistive Wheatstone bridge. In order to facilitate the building of this circuit without additional complexity, a fixed overheat ratio[5] was chosen. This ratio was set in the circuit by the three fixed resistors constructed close to the main sensor resistor. This provided the advantage of using the same material for all of the resistors, thereby minimizing resistor variations. Even if the different sensor resistors have varying resistances, the topological closeness of the other bridge resistors will help minimize the problems the circuitry may have in driving the sensor.

The resistance of the sensor also had to be changed going from Type I to Type II. The change was required for full compatibility with a 5v V_{dd} process. The earlier versions of the sensors used a bias of approximately 6v DC across them. This translates to an operating point for the Type I calculated by the following:

$$I_{sensor} = \frac{V_{sensor}}{R_{sensor}} \tag{7.1}$$

$$\begin{aligned} R_{cold} &= 2300\,\Omega \\ a_r &= 0.1 \\ I_{sensor} &= \frac{6\text{v}}{(1+a_r)R_{cold}} \\ &= 2.372\,\text{mA} \end{aligned} \tag{7.2}$$

$$\begin{aligned} P_{sensor} &= \frac{V_{sensor}^2}{R_{cold}} \\ &= 14.23\,\text{mW} \end{aligned} \tag{7.3}$$

[5]See Section 4.4.

To allow a lower operating voltage yet maintain the same sensor power, we needed to reduce the sensor resistance. The new resistance of 429 Ω was acceptable. The current versus voltage plot of the new sensor is given in Figure 7.5. Finally, the layout picture of the new sensor is shown in Figure 7.6.

The SPICE model of the sensor is the depicted in Figure 7.4 where the $F(O)$ function is written as:

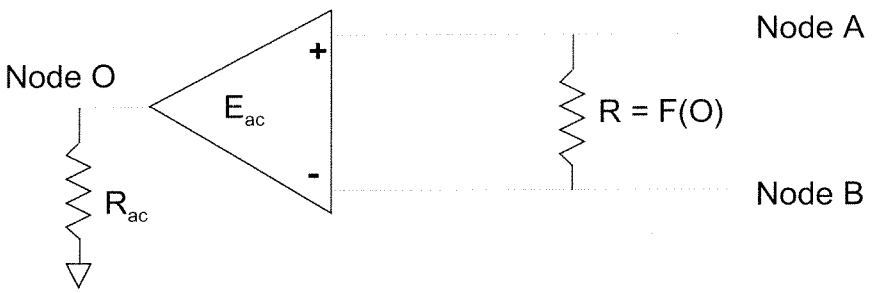

Figure 7.4: Schematic of the spice macromodel of the shear stress sensor. The function $F(O)$ is the empirically derived equation for the current versus voltage for the sensor and it is shown in Figure 7.5.

```
.subckt shear_sensor a b
$
$ xmgr derived parameters
.param a1 = 0.00228425625 a2 = 5.05918143e-5
.param a3 = -0.000116463538

$ convert v(a) - v(b) into a single variable v(o)
eac o gnd a b 1.0
rac o gnd 1000MEG

r1 a b r='.01 + v(o)/(a1*v(o)+a2*v(o)*v(o)+a3*v(o)*v(o)*v(o))'

.ends
```

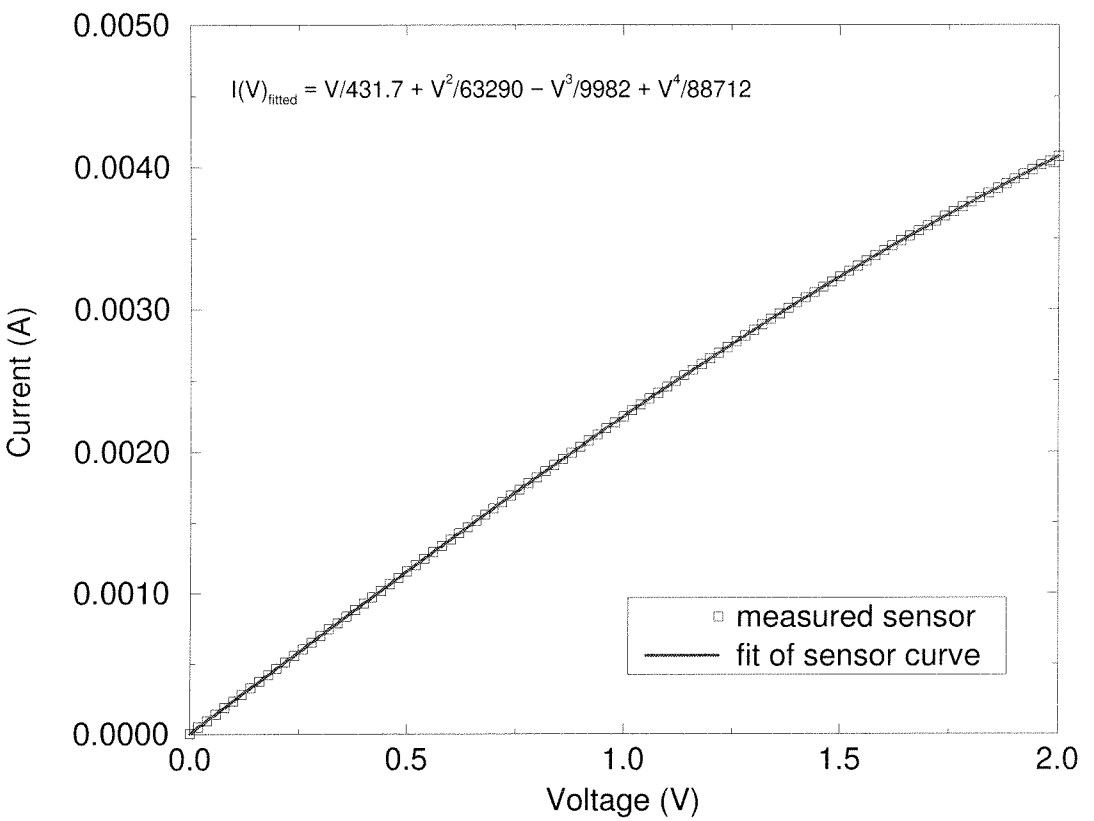

Figure 7.5: Current versus voltage curve for the low voltage compatible shear stress sensor (Type II).

On-Chip Constant Temperature Circuitry

The amplifier is a special design that allows the introduction of the imbalance parameter as a steerable current (see Figure 7.7). The opamp has a gain-bandwidth product of approximately 20 MHz (see Figure 7.8) and a DC gain of 66 dB (≈ 2000) (see Figure 7.9).

The whole CT system is depicted in Figure 7.10. The sensor in the schematic is actually represented by an equation which represents the fit of the data (Section 7.3.1 and Figure 7.5). Also included in the schematic is the test signal which helps to characterize the overall frequency response of the system. In the circuit shown here, the overall bandwidth is in excess of 10 KHz.

7.3.2 Integrated Actuator

The difficulty with building an actuator with a high enough actuation bandwidth has caused the design to evolve from a surface micromachinined version to a bulk micromachined actuator. This techniques allows the bulk silicon to be used to produce a beam with a higher actuation bandwidth[6].

Integrated μ-Actuator Driving Scheme

The fact that the μ-actuator will be driven at its resonant frequency presents some issues with respect to the driving circuitry. Without a clear mechanism to detect the actual movement of the μ-actuator, the feedback schemes available to ensure peak operation are limited. Rather, uniform fabrication results are assumed so that the spread among μ-actuator resonant frequencies is small. If this assumption is true, then the actuation scheme is to apply a sinewave at the peak of the resonant frequency distribution. The actuator is then turned on with the power switch being the output transistor of the control/detection circuitry.

[6]Due to bulk silicon's higher stiffness.

7.3.3 Dieplot for the Wafer-Scale Integration

Figure 7.11 and Figure 7.12 show the different layers of our integrated system having to do with CMOS and MEMS respectively.

7.4 Towards Total Wafer Scale Integration

The goal of the project to which this research has contributed is to build a complete system in order to test our hypotheses about sufficient control algorithms. There are a number of hurdles still to be overcome. The biggest hurdle is the μ-actuator bandwidth limitation. Another hurdle is the ability to integrate the micromachined components with a standard CMOS process. At the present time, Orbit Semiconductor cannot fabricate a complete wafer due to excessive sensor height. Although the sensor only protrudes 2μm above the surface, Orbit's 1μm thick photoresist cannot cover the surface adequately to create the patterns necessary for metallization. Clearly a solution to this problem must be found in order for fabrication to proceed.

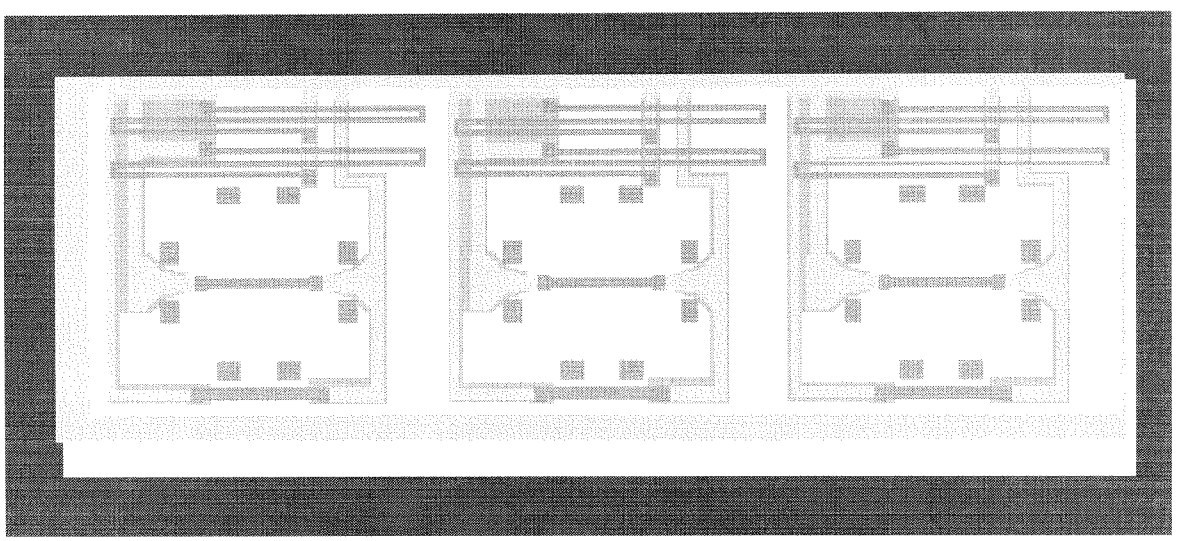

Figure 7.6: Sensor layout for integrated version showing all four resistors of the bridge circuit drawn in the same area to minimize matching problems.

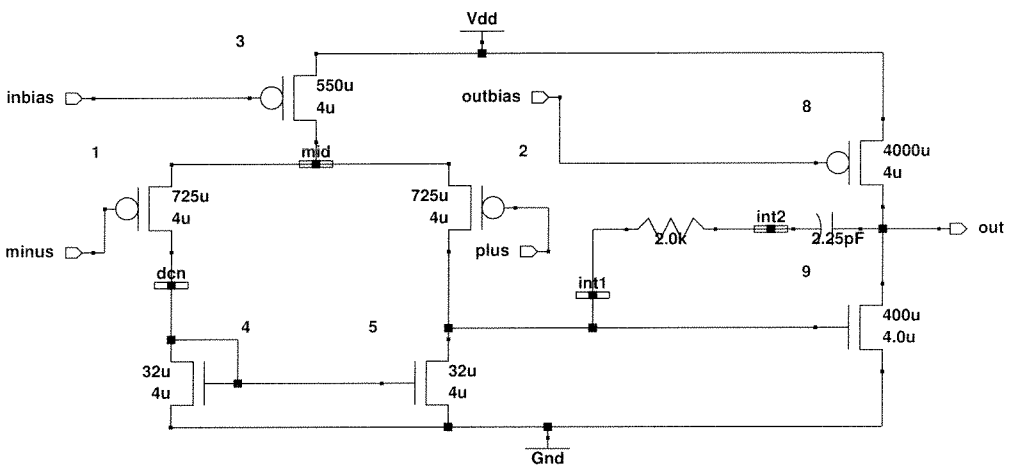

Figure 7.7: Schematic of the constant temperature operational amplifier.

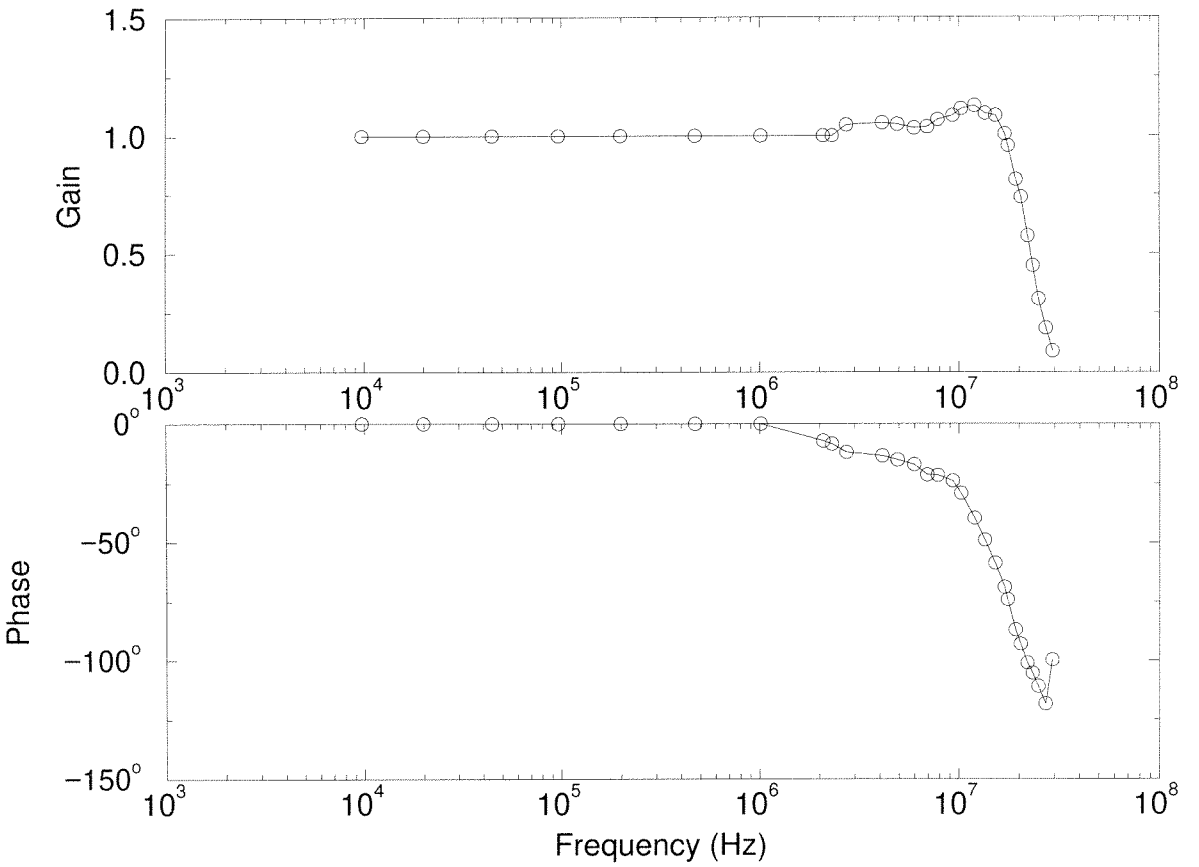

Figure 7.8: Graph of the constant temperature operational amplifier bandwidth and phase characteristics.

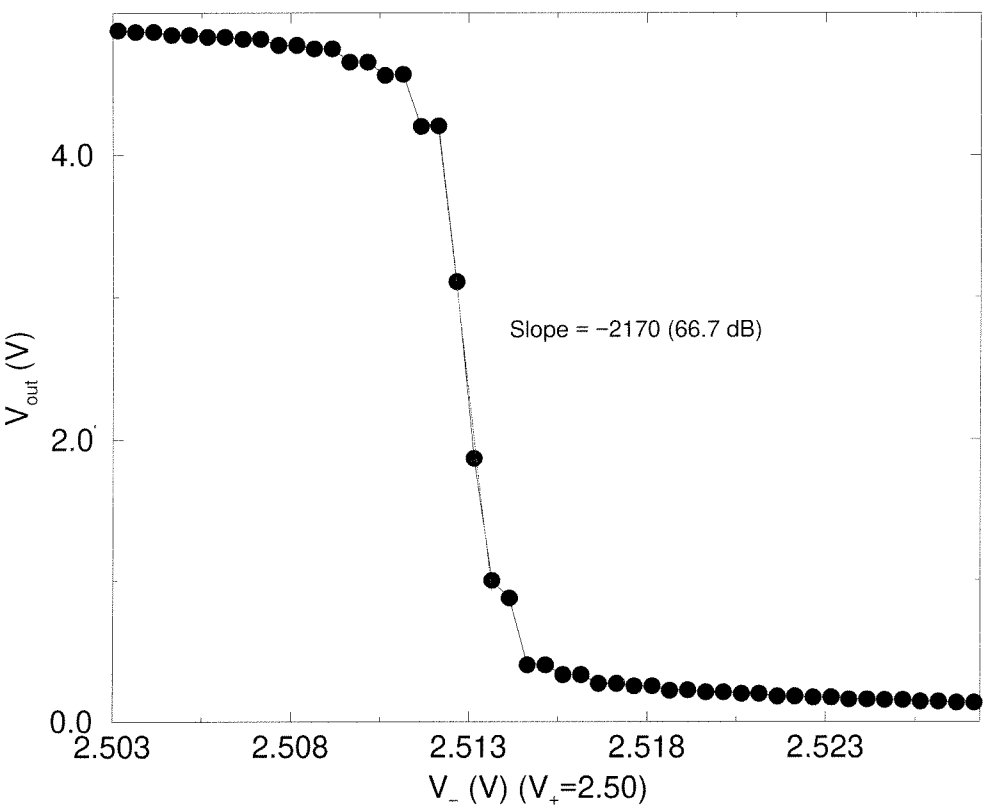

Figure 7.9: Plot of the constant temperature operational amplifier DC gain.

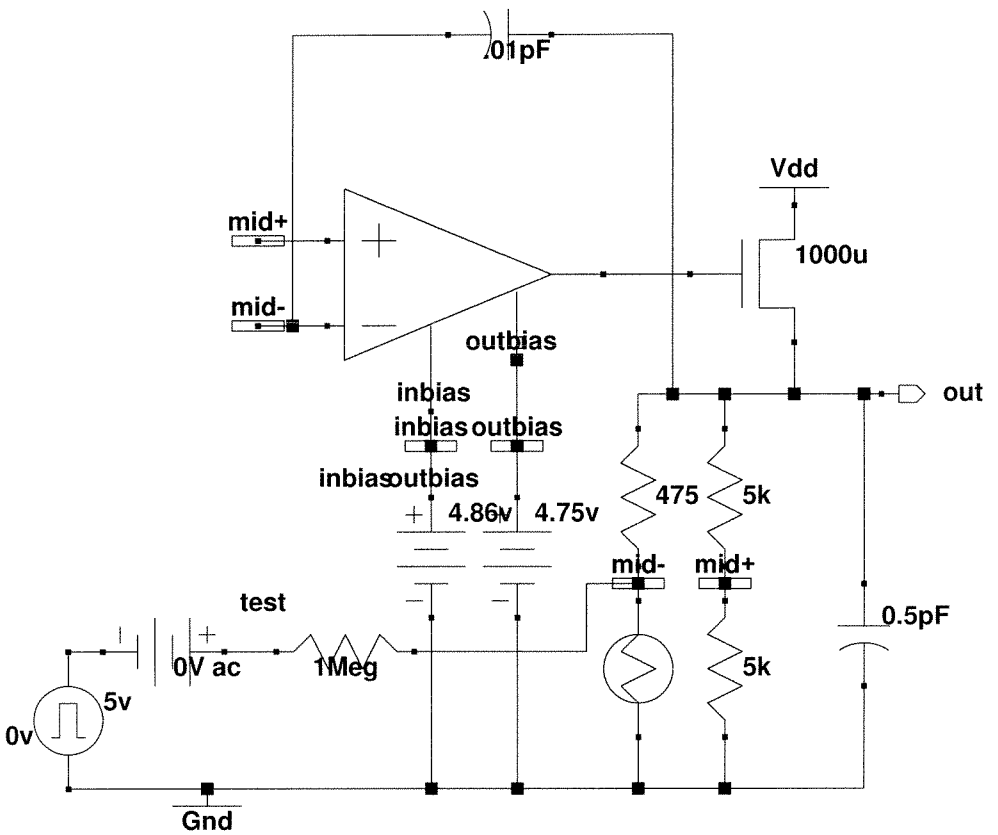

Figure 7.10: Schematic of the whole CT system which incorporates the constant temperature circuit.

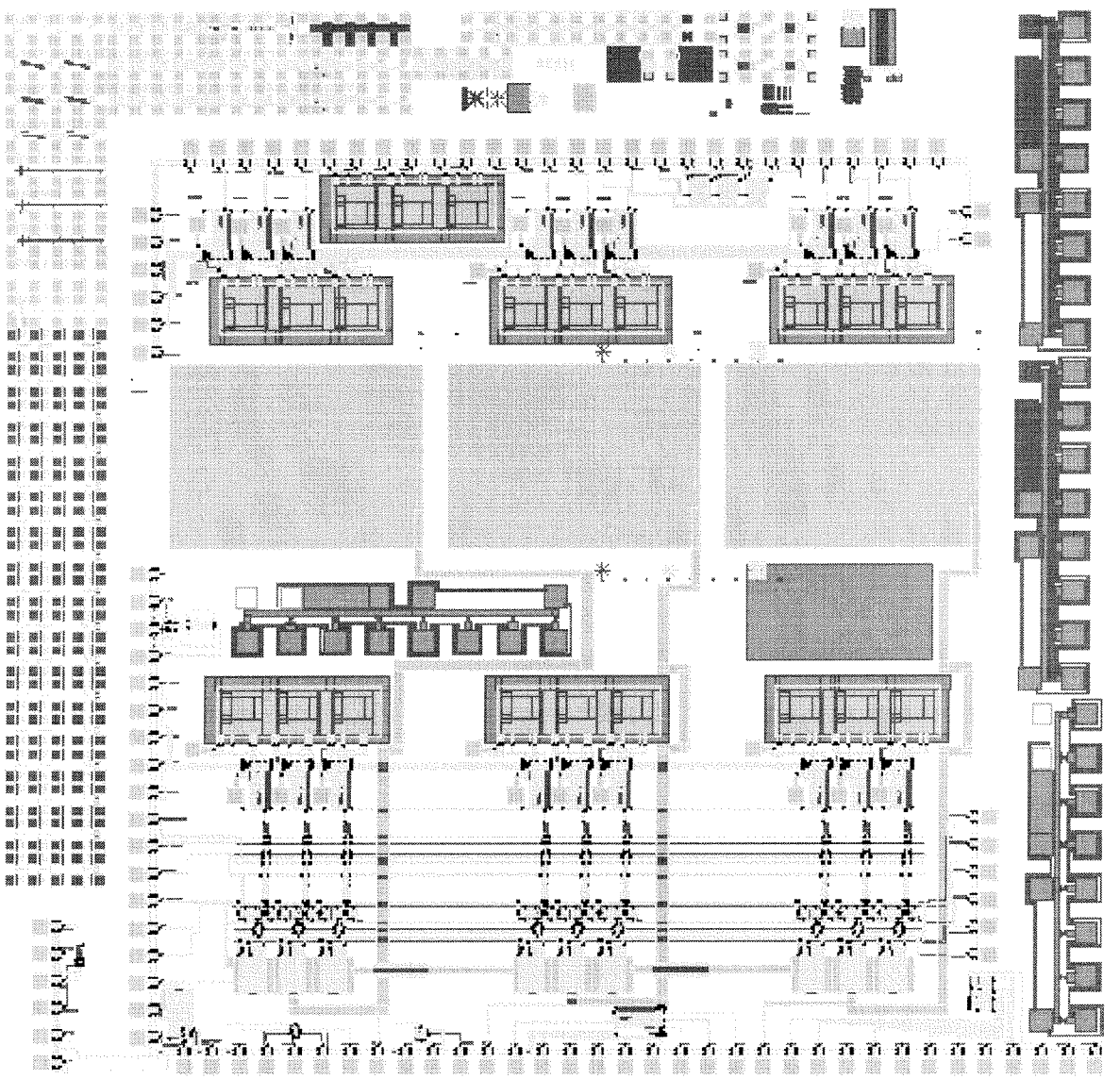

Figure 7.11: Picture of the CMOS layers on the die for wafer-scale integration.

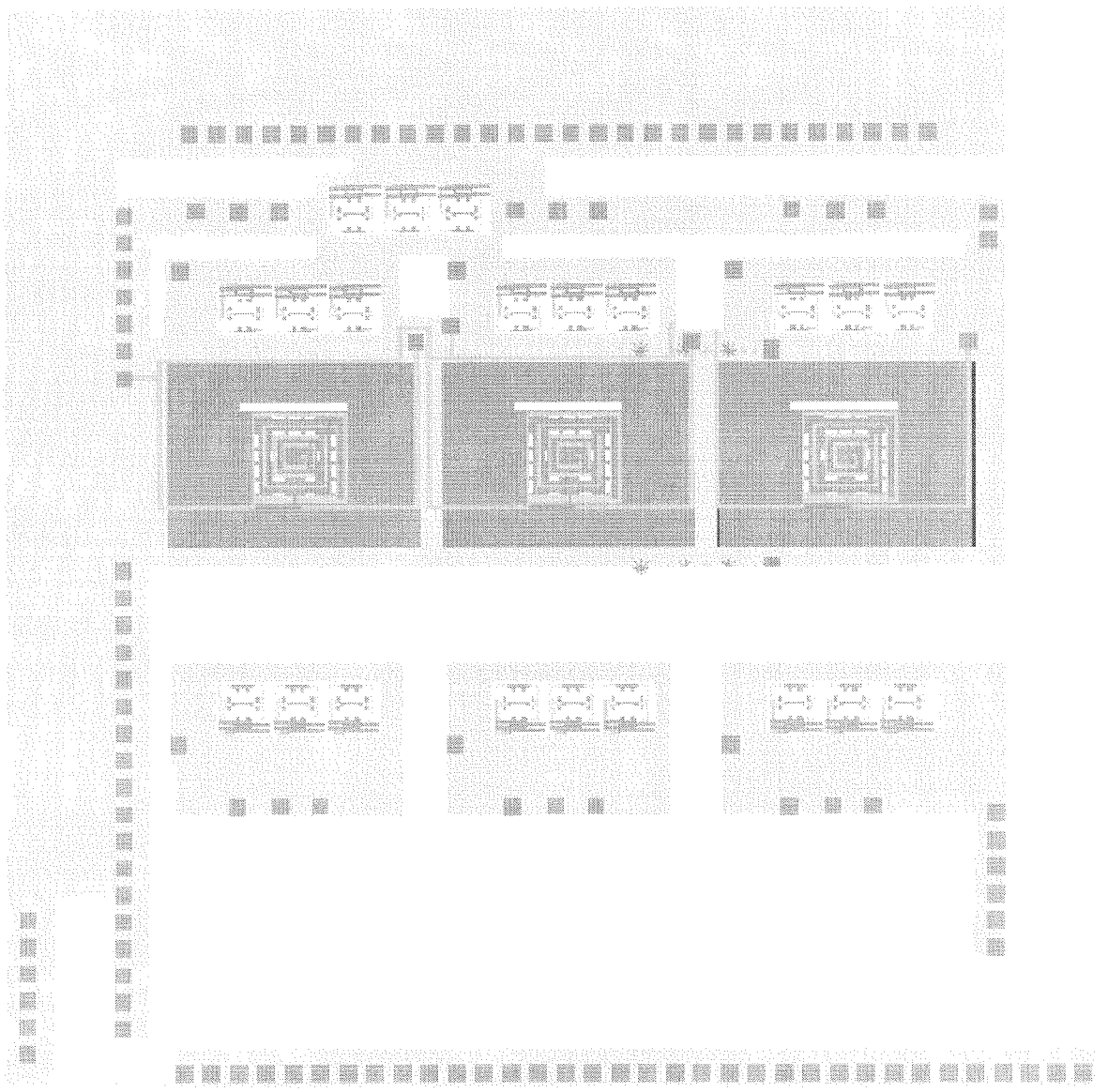

Figure 7.12: Picture of the MEMS layers on the die for wafer-scale integration.

Chapter 8 Conclusion

We have described an analog VLSI system that interfaces with microfabricated constant temperature shear stress sensors. This system detects regions of high shear stress and outputs a control signal to activate a microactuator. We are in the process of verifying the actual drag reduction by controlling microactuators in wind tunnel experiments. We are encouraged that our microscale approach, similar to ones biology may employ, can provide a very useful contribution to the problem of drag reduction.

8.1 Future Challenges

The processing system that we have built is a reasonable first pass at addressing the different issues of how to process the information. One drawback of the present system is that it is not very adaptable to changing conditions. For instance, should the velocity of the airflow change drastically, the circuit still would respond with the fixed topology, i.e. it only detects structures on the scale designed. While this is less than ideal, it is a reasonable solution for an initial approach to the problem.

Appendix A Constant Temperature Analysis

In this appendix, we analyze the constant temperature circuit with an eye towards the integrated version as well as noise calculations. We roughly follow the analysis found in [18, 19, 20, 7, 45].

A.1 Circuit

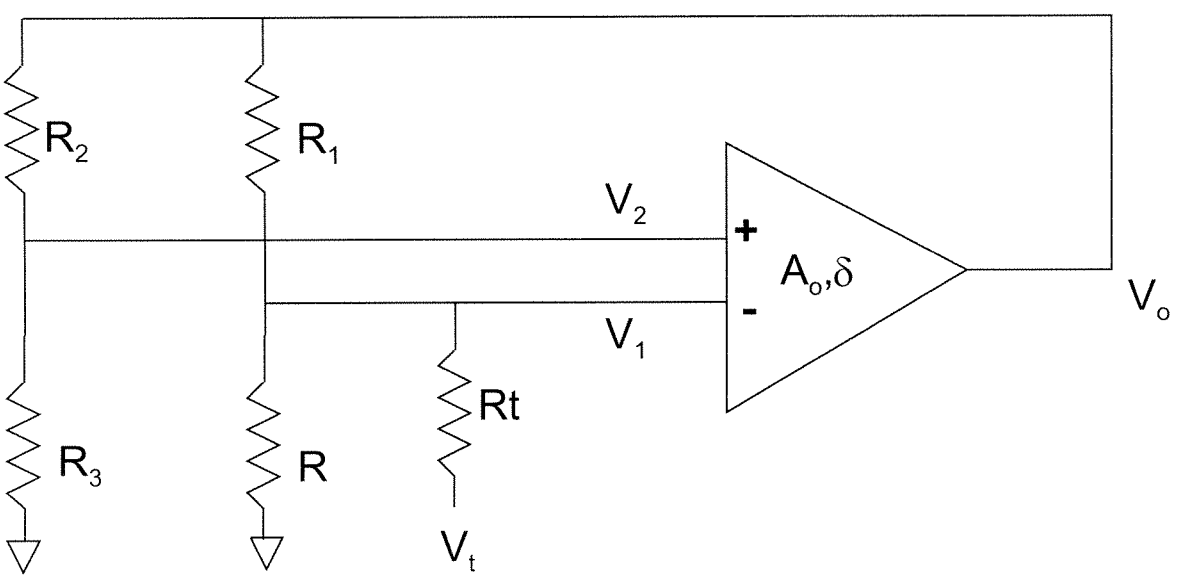

Figure A.1: Schematic of the constant temperature circuit we wish to analyze with the elements and nodes labeled.

We refer to Figure A.1 for the generic schematic of the circuit. We use the notation of the variable having both a quiescent as well as time-varying component (i.e., $V = \bar{V} + \tilde{v}$).

We start by denoting the sensor resistance as $R = \overline{R} + \tilde{R}$. $\overline{R}$ is the quiescent resistance of the sensor after it has been heated. This is related to the over heat

ratio (a_r) by the following equation:

$$\overline{R} = (1 + a_r)\, R_{cold} \tag{A.1}$$

Assuming that the input to the bridge is designated $V_o = \overline{V}_o + \tilde{V}_o$, we have the following relationship for the two arms of the bridge:

$$V_1 = V_o \frac{R}{R + R_1} + V_t \frac{RR_1}{R_t(R_1 + R)} \tag{A.2}$$

$$V_2 = V_o \frac{R_3}{R_2 + R_3} \tag{A.3}$$

$$= \frac{V_o}{2}$$

$$V = V_1 - V_2$$

$$= \frac{V_o}{2}\frac{R_1 - R}{R_1 + R} - \frac{V_t}{2}\frac{R}{R_t} \tag{A.4}$$

$$\tag{A.5}$$

where V_1 is the nodal voltage of the arm containing the shear stress sensor, V_2 is the middle node of the bridge arm with the fixed resistors and ΔV is the difference between the two. V_t is the electrical testing signal which is applied to the circuit via a resistor R_t.

We now write the equations for the closed loop circuit. We need to remember that there needs to be an imbalance in the opamp input in order for the circuit to find the correct operating point. With a resistive circuit the origin (meaning all nodes at zero voltage) is also a stable point that needs to be avoided.

The opamp has the following characteristic for two poles:

$$\frac{V_o}{V_i} = A_o \frac{1}{(1 - \frac{s}{p_1})(1 - \frac{s}{p_2})} \tag{A.6}$$

$$\text{rewriting} \tag{A.7}$$

$$A_o V_i + \delta = \frac{1}{p_1 p_2} s^2 V_o - \left(\frac{1}{p_1} + \frac{1}{p_2}\right) s V_o + V_o \tag{A.8}$$

where δ is the amplifier offset when $V_i = 0$v.

The by combining Equation A.4 and Equation A.8, the result is as follows:

$$\frac{1}{p_1 p_2} s^2 V_o - \left(\frac{1}{p_1} + \frac{1}{p_2}\right) s V_o + V_o \left(1 + \frac{A_o}{2} \frac{R - R_1}{R + R_1}\right) = -V_t \frac{A_o}{2} \frac{R}{R_t} + \delta \quad \text{(A.9)}$$

The equation which relates the sensor resistance to the flow velocity is given by [18]:

$$V_o^2 \frac{R}{(R + R_1)^2} - \frac{H(U)}{\alpha} (R - R_0) = \frac{c}{\alpha} s R \quad \text{(A.10)}$$

where $H(U)$ is the heat transfer function as a function of the velocity U and c is the thermal inertia of the sensor.

If we now expand Equation A.9 and Equation A.10 by writing the terms as both a constant term as well as a small-signal one and applying Taylor's expansion, $f(x) = f(a) + \acute{f}(a)(x - a)$, about the constant point, we obtain:

$$\frac{4R_1}{A_o} \frac{1}{p_1 p_2} s^2 \tilde{V}_o + \frac{4R_1}{A_o} \left(\frac{1}{p_1} + \frac{1}{p_2}\right) s \tilde{V}_o + \frac{4R_1}{A_o} \frac{\delta}{x} \tilde{V}_o + \tilde{R} = \frac{2R_1^2}{x R_t} V_t \quad \text{(A.11)}$$

$$\tilde{V}_o - \frac{x}{2(R_1 - R_0)} \left(\tilde{R} + \frac{c}{H(U)} s \tilde{R}\right) = \frac{V_o}{2H(U)} \frac{\partial H}{\partial U} \tilde{U} \quad \text{(A.12)}$$

where $x^2 = 4R_1 \frac{H(U)}{\alpha} (R_1 - R_0)$.

If we combine Equation A.11 and Equation A.12 by eliminating $\tilde{R}$, we arrive at:

$$\frac{2R_1 c}{A_o (R_1 - R_0) H(U)} \frac{1}{p_1 p_2} s^3 \tilde{V}_o +$$
$$\frac{2R_1 c}{A_o (R_1 - R_0) H(U)} \left(\frac{1}{p_1 p_2} \frac{H(U)}{c} + \left(\frac{1}{p_1} + \frac{1}{p_2}\right)\right) s^2 \tilde{V}_o +$$
$$\frac{2R_1 c}{A_o (R_1 - R_0) H(U)} \left(\left(\frac{1}{p_1} + \frac{1}{p_2}\right) \frac{H(U)}{c} + \frac{\delta}{x}\right) s \tilde{V}_o + \tilde{V}_o$$

$$= \frac{x}{2H(U)}\frac{\partial H}{\partial U}\left(1-\frac{2R_1}{R_1-R_0}\right)\tilde{U} \tag{A.13}$$

$$+\frac{2R_1^2}{R_t(R_1-R_0)}V_t+\frac{R_1 c}{R_t(R_1-R_0)H(U)}sV_t \tag{A.14}$$

What we are really interested in is the response of the system to the flow as well as to the electrical testing signal.

$$\text{flow transfer function} = \frac{\tilde{V_o}}{\tilde{U}} \tag{A.15}$$

$$= \frac{\frac{x}{2H(U)}\frac{\partial H}{\partial U}\left(1-\frac{2R_1}{R_1-R_0}\right)}{A_1s^3+A_2s^2+A_3s+1} \tag{A.16}$$

where

$$A_1 = \frac{2R_1c}{A_o(R_1-R_0)H(U)}\frac{1}{p_1p_2} \tag{A.17}$$

$$A_2 = \frac{2R_1c}{A_o(R_1-R_0)H(U)}\left(\frac{1}{p_1p_2}\frac{H(U)}{c}+\left(\frac{1}{p_1}+\frac{1}{p_2}\right)\right) \tag{A.18}$$

$$A_3 = \frac{2R_1c}{A_o(R_1-R_0)H(U)}\left(\left(\frac{1}{p_1}+\frac{1}{p_2}\right)\frac{H(U)}{c}+\frac{\delta}{x}\right) \tag{A.19}$$

$$\text{test transfer function} = \frac{\tilde{V_o}}{V_t} \tag{A.20}$$

$$= \frac{\frac{2R_1^2}{R_t(R_1-R_0)}+\frac{R_1c}{R_t(R_1-R_0)H(U)}s}{A_1s^3+A_2s^2+A_3s+1} \tag{A.21}$$

We use SPICE to give us the resulting poles of the closed-loop system. Since we model nearly all transistor parasitics in SPICE, the number of poles is larger than our hand calculation. The output of spice for the poles of the system, approximated by $A_1s^3+A_2s^2+A_3s+1$ is as follows:

poles (rad/sec)		poles (hertz)	
real	imag	real	imag
-58.8634k	0.	-9.3684k	0.
-165.4795x	0.	-26.3369x	0.
-196.4500x	-145.1815x	-31.2660x	-23.1064x
-196.4500x	145.1815x	-31.2660x	23.1064x
-347.2419x	-221.5152x	-55.2653x	-35.2552x
-347.2419x	221.5152x	-55.2653x	35.2552x
-2.8168g	0.	-448.3114x	0.
-7.2555g	0.	-1.1548g	0.

We observe that the first pole is approximately 9 kHz which is reasonable because we want to extract as much of the 10 kHz bandwidth as we can. Also the second pole magnitude is approximately 26 MHz which is greater than 2 orders of magnitude from the first pole. This gives the system a reasonable phase margin.

A.1.1 Noise

The thermal noise in a resistor is written as

$$v_{rms}^2 = 4kTR\Delta f \tag{A.22}$$

This is approximately equal to 4nV for 1 kΩ resistor when used with a 1 kHz bandwidth. For our Type I discrete sensors (@ 10 kHz bandwidth), we calculate this value as:

$$v_{rms}^2 = 4kT(2.3K\Omega)(10kHz) \tag{A.23}$$

$$v_{rms} = 19.2nV \tag{A.24}$$

For the Type II integrated sensor with lower resistance, we calculate this value to be

$$v_{rms} = 8.3nV \tag{A.25}$$

Both of these values for the rms noise voltage are relatively low compared to all of the other sources of noise in the circuit. And it also means that the signal to noise ratio can be rather high.

The noise of a constant temperature hot-film system has been derived by Freymuth [20]:

$$|\tilde{V}_n| = \frac{2R}{R - R_0}\left|1 + M\sqrt{\frac{j\omega}{\omega_c}}\frac{\acute{I}_0}{I_0}\right| A_n \tag{A.26}$$

where M is the thermal time constant of the sensor, ω_c is the critical frequency, I_0 is the modified Bessel function of zeroth order and A_n is the input referred amplifier noise voltage.

If we evaluate this equation using values from our system, we obtain the following values:

$$|\tilde{V}_n| = \frac{2(1+\alpha_r)}{\alpha_r}\left|1 + M\sqrt{\frac{j\omega}{\omega_c}}\frac{\acute{I}_0}{I_0}\right| A_n \tag{A.27}$$

$$\approx 22A_nF(\omega) \tag{A.28}$$

where $F(\omega)$ is a simplified form of the above expression. If we evaluate the noise at various frequencies relative to M and ω_c, we calculate the following:

$$\omega << M \qquad |\tilde{V}_n| = 22A_n \tag{A.29}$$

$$M < \omega < \omega_c \qquad |\tilde{V}_n| = 22A_nM\omega \tag{A.30}$$

$$\omega_c < \omega \qquad |\tilde{V}_n| = 22A_nM\sqrt{\omega\omega_c} \tag{A.31}$$

The input-referred noise of the opamp (A_n) is calculated using SPICE to be $A_n = 1.79\frac{\mu v}{\sqrt{\text{Hz}}}$. Hence the opamp noise when we use it to 10 kHz is:

$$A_n^2 = 3.210^{-12}\frac{\text{v}^2}{\text{Hz}} * 10\text{kHz} \tag{A.32}$$

$$A_n = 179\mu\text{v} \tag{A.33}$$

And when we consider the important case of $M < \omega < \omega_c$ for our integrated CT circuit, we observe that

$$|\tilde{V}_n| = 22\,179\mu\text{v}\,\frac{1}{1750\,2\pi}\,\omega \tag{A.34}$$

$$\tilde{V}_n^2 = 15.6\,\frac{\mu\text{v}}{\text{Hz}}\,2\pi\omega \tag{A.35}$$

$$\tilde{V}_n = 3.95\text{mv} \tag{A.36}$$

This value of 3.95 mv for $\tilde{V}_n$ at a frequency of 1750 Hz is consistent with what we measure from the simulations of the constant temperature circuit.

Appendix B Process Flow for Berkeley Integration

This is the detailed process flow for building a fully integrated system using the Berkeley 2.0μm CMOS process. The sensors and actuators are the surface micro-machined variety.

```
1) start with p-type Si <100> 8-12 ohm/sq
2) grow SiO2 100nm
3) define well [N Well] {CWN}
4) implant P 5e12/cm2 @ 150KeV
5) drive-in well to 3.4um
6) grow SiO2 30nm (pad ox)
7) deposit Si3N4 100nm
8) define active area (PR1) [Active] {CAA}
9) etch Si3N4 (where LOCOS)
10) define p-field (PR2) [??] {PFLD}
11) implant B 1.5e13/cm2 @ 70 KeV
12) grow LOCOS 650-700nm
13) remove nitride (from 7)
14) remove oxide (from 6 - pad ox)
15) grow sac. oxide 20nm
16) implant B 1.7e12/cm2 @ 30 KeV
17) remove sac. oxide (from 15)
18) grow gate oxide 30nm
19) deposit gate poly 450nm [Poly] {CPG}
20) define gate poly
```

21) etch gate poly (from 19)

22) deposit cap. oxide 80nm

23) deposit cap. poly 450nm [Poly2] {CEL}

24) define cap. poly

25) etch cap. poly

26) define N+ s/d

27) implant As 5e15/cm2 @ 160 KeV

28) anneal N+ s/d

29) deposit PSG 700nm

*** back to Caltech at this point ***

30) etch PSG (from 29) [fberk_psg] {FEP} <bHF 2000-4000A/min.>

31) deposit PSG 400nm
<@450C SiH4: 20 sccm, O2: 62 sccm, PH3: 10 sccm, 75A/min.>

32) deposit poly 1um <@620C SiH4: 80 sccm, 140 A/min.>

33) pattern poly [tpoly] {TNP}
<in PEIIA 200W, CF4/O2=200mT/120mT, 1500 A/min.>

34) pattern PSG (from 31) [fcaltech_psg] {NG} <bHf>

35) deposit Si3N4 1.2um <@835C NH3: 15.6 sccm, DCS: 64.7 sccm>

36a) pattern Si3N4 for etch holes [fhole_etch] {NS}
<in RIE 600W, SF6/O2=65 sccm/15 sccm 700 A/min.>

36b) etch cavity w/49% HF (25 min.)

37) deposit LTO 1um <at 450C SiH4: 42.8, O2: 62 sccm>

38) deposit Si3N4 400nm <@835C, 75 min.>

39) pattern Si3N4 [fseal_pad_nitride] {NB}
<in PEIIA 200W, CF4/O2=200mT/120mT, 1000A/min.>

40) pattern LTO (from 37) [fseal_pad_lto] {FLTO} <bHF 1500 A/min.>

41) deposit poly 600nm at 560C (SH4: 80sccm, 160 min.)

42a) dope poly B 1e16 @80kev

42b) anneal poly at 950C for 1 hour.

43) pattern poly (from 41) [fpoly] {NP}
    <in RIE 600W, SF6/O2= 65 sccm/15 sccm, 1000 A/min.>

44) deposit Si3N4 200nm <@835C, 40 min.>

45) pattern Si3N4 [fcontact] {NC} <in RIE>

46) etch LTO (from 37) <bHF>

47) pattern Si3N4 (from 35) [fnitride] {FN}
    <in PEIIA 200W, CF4/O2=200 mT/120 mT, 1000A/min.>

**** return to Berkeley ******

48) sputter Al 600nm

49) pattern Al (Lam/plasma) [Metal1] {CMF}

**** back to Caltech ******

50) pattern Al (wet etch) [fberk_metal] {NM}

51) pattern Si3N4 (from 35) [tnitride_etch] {TNC}
    <etch w/CF4 @ 200W, 300mT, rate 1000 A/min>

52) deposit LTO 2um
    < 160 min.>

53) pattern LTO [tlto_etch] {TNG}
    <etch w/bHF, rate 4000 A/min>

54) deposit Cr/Au 400nm

55) pattern Cr/Au [tmetal1+] {TNM}
    <Au etch - about 1200 A/min>
    <Cr mask etch - fast>

56) deposit LTO 500nm

57) pattern LTO [tlto2_etch] {TLYL}
    <100 A Cr, 1000 A Cu>

58) deposit Cr/Au 400nm

59) pattern Cr/Au [tmetal2+] {TNMS}

60) deposit seed layer (PR)

61) pattern seed layer [tplating] {TNS}

62) plate NiFe

    <cal. for rate 5um/hr>

63) remove PR

    <acetone + ASCI PR stripper>

64) protect NiFe [tcr_protect] {TLYR}

65) remove seed layer

    <Cu etch - 10% H2O2, 10% acetic acid>

    <Cr etch - Cr-7 etch>

66) etch w/TMAH

    <TMAH undercut rate 1.4um/min>

Bibliography

[1] N. Aubry, P. Holmes, J.L. Lumley, and E. Stone. Behavior of coherent structures in the wall region by dynamical systems. In S.J. Kline and N.H. Afgan, editors, *Near-wall Turbulence : 1988 Zoran Zaric Memorial Conference*, pages 672–691. Hemisphere Publishing Corp., New York, NY, 1988.

[2] D. Babcock, C. Lee, B. Gupta, J. Kim, and R. Goodman. Active drag reduction using neural networks. In *Proceedings of the International Workshop on Neural Networks for Identification, Control, Robotics, and Signal/Image Processing*, Venice, Italy, August 1996.

[3] P.R. Bandyopadhyah. Reynolds number dependence of the freestream turbulence effects on turbulent boundary layers. *AIAA Journal*, 30(7):1910–1912, July 1992.

[4] A. Baron and M. Quadrio. Influence of riblets on a turbulent boundary layer. *International Journal of Heat and Fluid Flow*, 14(3):223–230, September 1993.

[5] D. W. Bechert and M. Bartenwerfer. The viscous flow on surfaces with longitudinal ribs. *Journal of Fluid Mechanics*, 206:105–129, 1989.

[6] D. W. Bechert, G. Hoppe, and W.-E. Reif. On the drag reduction of the shark skin. *Proceedings of the AIAA Shear Flow Control Conference*, AIAA(85-0546):1–18, March 1985.

[7] R.F. Blackwelder. Hot-wire and hot-film anemometers. In R.J. Emrich, editor, *Fluid Mechanics*, volume 18 of *Methods of Experimental Physics*, pages 259–314. Academic Press, New York, 1981.

[8] R.F. Blackwelder and R.E. Kaplan. On the wall structure of the turbulent boundary layer. *Journal of Fluid Mechanics*, 76:89–112, 1976.

[9] H.H. Bruun. *Hot-Wire Anemometry: Principles and Signal Analysis.* Oxford University Press, 1995.

[10] D.M. Bushnell and K.J. Moore. Drag reduction in nature. *Annual Review Fluid Mechanics*, 23:65–79, 1991.

[11] H.J. Catrakis and P.E. Dimotakis. Scale distributions and fractal dimensions in turbulence. *Physical Review Letters*, 77(18):3795–3798, October 1996.

[12] D.R. Chapman and G.D. Kuhn. The limiting behavior of turbulence near a wall. *Journal of Fluid Mechanics*, 170:265–292, 1986.

[13] H. Choi, P. Moin, and J. Kim. Turbulent drag reduction: Studies of feedback control and flow over riblets. Technical Report TF-55, Stanford University, Stanford, Ca., 94305, September 1992.

[14] H. Choi, P. Moin, and J. Kim. Active turbulence control for drag reduction in wall-bounded flows. *Journal of Fluid Mechanics*, 262:75–110, 1994.

[15] Tobias Delbrück. *Investigations of Analog VLSI Visual Transduction and Motion Processing.* Ph.D. thesis, California Institute of Technology, 1993.

[16] G. Dingerkus and R.J. Koestler. Application of scanning electron microscopy to the study of shark dermal denticles. *Scanning Electron Microscopy*, II:513–519, 1986.

[17] Encyclopaedia britannica, 1996.

[18] P. Freymuth. Frequency response and electronic testing for constant-temperature hot-wire anemometers. *Journal of Physics E: Scientific Instruments*, 10:705–710, 1977.

[19] P. Freymuth. Further investigation of the non-linear theory for constant-temperature hot-wire anemometers. *Journal of Physics E: Scientific Instruments*, 10:710–713, 1977.

[20] P. Freymuth. A comparative study of the signal-to-noise ratio for hot-film and hot-wire anemometers. *Journal of Physics E: Scientific Instruments*, 11:915–918, 1978.

[21] M. Gad–el–Hak. Drag reduction: An overview. In *Drag Reduction Techniques*.

[22] M. Gad–el–Hak. The art and science of flow control. In M. Gad–el–Hak, editor, *Frontiers in Experimental Fluid Mechanics*, volume 46 of *Lecture Notes in Engineering*, pages 211–290. Springer-Verlag, Berlin, 1989.

[23] M. Gad–el–Hak. Interactive control of turbulent boundary layers: A futuristic overview. *AIAA Journal*, 32(9):1753–1765, September 1994.

[24] Gad–el–Hak, M. and P.R. Bandyopadhyay. Field versus laboratory turbulent boundary layers. *AIAA Journal*, 33(2):361–364, February 1995.

[25] Gad–el–Hak, M. and R.F. Blackwelder. Selective suction for controlling bursting events in a boundary layer. *AIAA Journal*, 27(3):308–314, 1988.

[26] A. Hussain and W.C. Reynolds. The mechanics of a perturbation wave in turbulent shear flow. Scientific Report 1655TR, AFOSR, 1970.

[27] S.A Jacobson and W.C. Reynolds. Active control of boundary layer wall shear stress using self-learning neural networks. *Proceedings of the AIAA Shear Flow Control Conference*, AIAA(93-3273):1–12, July 1993.

[28] F. Jiang, Y.-C. Tai, C.-M. Ho, R. Karan, and M. Garstenauer. Theoretical and experimental studies of the micromachined hot-wire anemometers. *Technical Digest of the 1994 IEDM Conference*, pages 139–142, 1994.

[29] F. Jiang, Y.-C. Tai, C.-M. Ho, and W.J. Li. A micromachined polysilicon hot-wire anemometer. *Technical Digest of the Solid-State Sensor and Actuator Workshop*, pages 264–267, 1994.

[30] F. Jiang, Y.-C. Tai, J.B. Huang, and C.-M. Ho. Polysilicon structures for shear stress sensors. *Digest of the IEEE TENCON 1995*, pages x–y, 1995.

[31] Sanjay S. Joshi. *A Systems Theory Approach to the Control of Plane Poiseuille Flow.* Ph.D. thesis, University of California Los Angeles, 1996.

[32] Y. Katsuki and T. Hashimoto. Shark pit organs: Enhancement of mechanoreceptivity by potassium ions. *Science*, 166:1287–1289, 1969.

[33] Y. Katsuki, K. Yanagisawa, A.L. Tester, and J.I. Kendall. Shark pit organs: Response to chemicals. *Science*, 163:405–407, 1969.

[34] Douglas Kerns. *Experiments in very large-scale analog computation.* Ph.D. thesis, California Institute of Technology, 1993.

[35] J. Kim, P. Moin, and Moser. Turbulence statistics in fully-developed channel flow at low reynolds-number. *Journal of Fluid Mechanics*, 162:339–363, 1987.

[36] J.C. Klewicki and R.E. Falco. On accurately measuring statistics associated with small-scale structure in turbulent boundary layers using hot-wire probes. *Journal of Fluid Mechanics*, 219:119–142, 1990.

[37] C. Lee, J. Kim, D. Babcock, and R. Goodman. Application of neural networks to turbulence control for drag reduction. In *Proceedings of the Flow Control Workshop*, Cargese, Corsica, France, 1996.

[38] C. Lee, J. Kim, D. Babcock, and R. Goodman. Application of neural network to turbulence control for drag reduction. *Physics of Fluids*, 1997. to appear.

[39] C. Liu, Y.-C. Tai, J.B. Huang, and C.-M. Ho. Surface micromachined thermal shear stress sensor. *ASME Application of Microfabrication to Fluid Mechanics*, pages 9–15, 1994.

[40] Chang Liu. *Silicon Micromachined Sensors and Actuators for Fluid Mechanics Applications.* Ph.D. thesis, California Institute of Technology, 1995.

[41] L. Löfdahl, G. Stemme, and B. Johansson. Silicon based flow sensors used for mean velocity and turbulence measurements. *Experiments in Fluids*, 12:270–276, 1992.

[42] Carver Mead. *Analog VLSI and Neural Systems.* VLSI Systems Series. Addison-Wesley, 1989.

[43] R.Y. Myose and R.F. Blackwelder. Control of streamwise vortices using selective suction. *AIAA Journal*, 33(6):1076–1080, 1995.

[44] Y. Nagano and T. Tsuji. Recent developments in hot- and cold-wire techniques for measurements in turbulent shear flows near walls. *Experimental Thermal and Fluid Science*, 9:94–110, 1994.

[45] A.E. Perry. *Hot-wire Anemometry.* Clarendon Press, 1982.

[46] W. Raschi and C. Tabit. Functional aspects of placoid scales: A review and update. *Australian Journal of Marine and Freshwater Research*, 43:123–147, 1992.

[47] W.E. Reif. Morphogenesis and function of the squamation in sharks. *Neues Jahrbuch für Geologie und Paläontologie Abhandlungen*, 164:172–183, 1982.

[48] W.E. Reif and A. Dinkelacker. Hydrodynamics of the squamation in fast swimming sharks. *Neues Jahrbuch für Geologie und Paläontologie Abhandlungen*, 164:184–187, 1982.

[49] M.A. Schmidt, R.T. Howe, S.D. Senturia, and J.H. Haritonidis. Design and calibration of a microfabricated floating-element shear-stress sensor. *IEEE Transactions on Electron Devices*, 35(6):750–757, June 1988.

[50] Martin Schmidt. *Microsensors for the measurement of shear forces in turbulent boundary layers.* Ph.D. thesis, Massachusetts Institute of Technology, 1988.

[51] Y.P. Tang and D.G. Clark. On near-wall turbulence generating events in a turbulent boundary layer on a riblet surface. In K.K. Prasad, editor, *Further Developments in Turbulence Management*, number 50 in Applied Scientific Research, pages 215–232. Kluwer Academic Publishers, Netherlands, 1993.

[52] A.L. Tester and G.J. Nelson. Free neuromasts (pit organs) in sharks. In P.W. Gilbert, R.F. Mathewson, and D.P. Randall, editors, *Sharks, Skates, and Rays*, chapter 34, pages 503–532. Johns Hopkins Press, Baltimore, MD, 1967.

[53] S. Tung, 1996. personal communication.

[54] Van der Hoeven, J.G. and D.W. Bechert. Experiments with a 1:4.2 model of a commuter aircraft with riblets in a large wind tunnel. In K.-S. Choi, editor, *Recent Developments in Turbulence Management*, pages 3–24. Kluwer Academic Publishers, Netherlands, 1991.

[55] S.A. Wainwright, F. Vosburgh, and J.H. Hebrank. Shark skin: Function in locomotion. *Science*, 202:747–749, 1978.

[56] C.E. Wark, A.M. Naguib, and H.M. Nagib. Effect of plate manipulators on coherent structures in a turbulent boundary layer. *AIAA Journal*, 28(11):1877–1884, November 1990.

[57] J.M. Younce. Mirrors on a chip. *IEEE Spectrum*, pages 27–31, November 1993.